

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2005, 23:53:58 ; Search time 31202 Seconds  
(without alignments)  
11774.484 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacaatcggtataaacc.....tattaaatctgcarctgcr 7582

Scoring table: IDENTITY NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7489.6	98.8	7582	6	BD196245	BD196245 Endogenet
2	7489.6	98.8	7582	6	BD267487	BD267487 Endogenet
3	7489.6	98.8	7582	6	AX000966	AX000966 Sequence
4	7489.6	98.8	7582	6	AX027480	AX027480 Sequence
5	6448	85.0	10222	9	AY101582	AY101582 Homo sapi
6	6448	85.0	10222	9	AY101585	AY101585 Homo sapi
7	6448	85.0	56093	9	HSAC000064	AX329572 Sequence
8	6448	85.0	149194	9	AC007566	AC000064 Human BAC
9	6448	85.0	10222	9	AY101583	AC007566 Homo sapi
10	6446.4	85.0	10222	9	AY101584	AY101583 Homo sapi
11	6440	84.9	10222	9	BD221808	AY101584 Homo sapi
12	6436	84.9	10499	6	AX007980	BD221808 Nucleic s
13	6436	84.9	10499	6	AY101586	AX007980 Sequence
14	6346.4	83.7	10229	9	AY101587	AY101586 Pan trogl
15	6344.8	83.7	10229	9	AC145964	AY101587 Pan trogl
16	6344.8	83.7	184675	9	AY101588	AC145964 Pan trogl
17	6269.6	82.7	10230	9	AY101589	AY101588 Gorilla g
18	6269.6	82.7	10230	9	AY101590	AY101589 Gorilla g
19	6037.8	79.6	10122	9	AY101590	AY101590 Pongo pyg

20	6025.8	79.5	10124	9	AY101591	Pongo pyg
21	5956.8	78.6	10246	9	AY101593	Hylobates
22	5950.4	78.5	10248	9	AY101592	Hylobates
C	5824.2	76.8	158033	9	AC018926	Hylobates
23	5492	72.4	105989	9	AC008121	Homo sapi
C	5492	72.4	110000	2	AC009727_2	Homo sapi
C	4755.6	62.7	187321	2	AC092510	Continuation (3 of
26	4302.2	56.7	114621	2	AC005187	Papio anu
C	4295.8	56.7	46575	2	AC080036	Homo sapi
C	4162	54.9	111140	9	AL592310	Homo sapi
C	3944.6	52.0	183680	9	AL592310	Human DNA
C	3779.8	49.9	73070	9	AC098859	Homo sapi
31	3731.6	49.2	173788	9	AL590143	Homo sapi
32	3731.6	49.2	173788	9	AC022555	Human DNA
33	3731.6	49.2	186540	9	AC090341	Homo sapi
34	3731.6	49.2	187837	9	AC079065	Homo sapi
C	3716.6	49.0	160883	2	AL158814	Homo sapi
35	3537.2	46.7	137947	9	HS4514	Homo sapi
C	3537.2	45.4	153444	2	AC040948	Homo sapi
C	3442.8	45.4	205035	9	CNS00008	Human chr
C	3442.8	45.4	205035	9	CNS00008	Human DNA
38	3438.2	45.3	101846	9	HS19701	Homo sapi
39	3438.2	45.3	224629	2	AC139260	Homo sapi
40	3438.2	45.3	224629	2	AC139260	Homo sapi
41	3420.6	45.1	68693	9	AC108746	Homo sapi
C	3367.4	44.4	118312	9	AC055713	Homo sapi
C	3367.4	44.4	197549	2	AC145423	Homo sapi
43	3300.6	43.5	169250	9	AP002812	Homo sapi
C	3300.6	43.5	169250	9	AP002812	Homo sapi
44	3290.2	43.4	134412	2	AP002377	Homo sapi
C	3290.2	43.4	134412	2	AP002377	Homo sapi

ALIGNMENTS

RESULT 1  
BD196245  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BD196245 7582 bp DNA linear PAT 17-JUL-2003  
Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.  
BD196245 GI:33006015  
JP 2002512530-A/11.  
unidentified  
unclassified  
1 (bases 1 to 7582)  
Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.  
Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders  
Patent: JP 2002512530-A 11 23-APR-2002;  
BIO MERIEUX  
OS Unidentified  
PN JP 2002512530-A/11  
PD 23-APR-2002  
PF 06-JUL-1998 JP 1999508244  
PR 07-JUL-1997 FR 97/08815  
PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND.  
PI FRANCOIS MALLET  
PC C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC  
Strandedness: Single;  
CC Topology: Linear;  
CC Endogenetic retroviral sequences, associated with autoimmune diseases  
CC and/or with pregnancy disorders  
FH Key Location/Qualifiers  
FT source 1. 7582  
/organism='Unidentified'.  
Location/Qualifiers  
1. 7582  
/organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'

FEATURES

source  
ORIGIN

Query Match 98.8% ; Score 7489.6 ; DB 6 ; Length 7582;

Best Local Similarity 100.0%; Pred. No. 0;		Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CAACAAATCGGGATATAAACCCAGGCAATTCGAGCTGSCAACAGCAGCCCCCTTTGGGTCC	60
DB	1	CAACAAATCGGGATATAAACCCAGGCAATTCGAGCTGSCAACAGCAGCCCCCTTTGGGTCC	60
QY	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTTGCAACTGCA	120
DB	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTTGCAACTGCA	120
QY	121	CTCTTCTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC	180
DB	121	CTCTTCTGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC	180
QY	181	TGTTTGGCAACCCGACGCTGCGCTGACTCCCATCTGCTCTGATCTCTGAGGCTGTC	240
DB	181	TGTTTGGCAACCCGACGCTGCGCTGACTCCCATCTGCTCTGATCTCTGAGGCTGTC	240
QY	241	CGCTGTGCTCTGATCCAGCGAGCGCCCATTCGCGCTCCCAATTTGGGCTTAAAGGCTTGC	300
DB	241	CGCTGTGCTCTGATCCAGCGAGCGCCCATTCGCGCTCCCAATTTGGGCTTAAAGGCTTGC	300
QY	301	CATTGTNCTGCAACGCTTAAGTGCCTGGTGTGTTCTAATTTGAGCTGAACTANTCACT	360
DB	301	CATTGTNCTGCAACGCTTAAGTGCCTGGTGTGTTCTAATTTGAGCTGAACTANTCACT	360
QY	361	GGGTTCCATGGTCTCTTCTGTGACCCAGGCTTCTAATAAATTAACATACTTACCACA	420
DB	361	GGGTTCCATGGTCTCTTCTGTGACCCAGGCTTCTAATAAATTAACATACTTACCACA	420
QY	421	TGGCCCCAAGATTCATTTCTTGGAAATCCGTGAGSCAAACGAACTCCAGGTTCAGAGAAATAC	480
DB	421	TGGCCCCAAGATTCATTTCTTGGAAATCCGTGAGSCAAACGAACTCCAGGTTCAGAGAAATAC	480
QY	481	GARGCTTGGCCACCATCTTGGAAAGCGGCTGCTACCTTCTTGGAAAGTGGTTCACCAATC	540
DB	481	GARGCTTGGCCACCATCTTGGAAAGCGGCTGCTACCTTCTTGGAAAGTGGTTCACCAATC	540
QY	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTRACATTTTGGCBACCMBSACGGACATCC	600
DB	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTRACATTTTGGCBACCMBSACGGACATCC	600
QY	601	MAAGTCAATGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGAGTATTTCTGGARAAT	660
DB	601	MAAGTCAATGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGAGTATTTCTGGARAAT	660
QY	661	TGGGAMCAATTTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTGC	720
DB	661	TGGGAMCAATTTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTGC	720
QY	721	GCCTGGACCTCTGAGGGAAGTATTAATTAATACCAATCTTACAGTAGACTCTTTTG	780
DB	721	GCCTGGACCTCTGAGGGAAGTATTAATTAATACCAATCTTACAGTAGACTCTTTTG	780
QY	781	TAGAAAGGCAATGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGACAA	840
DB	781	TAGAAAGGCAATGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGACAA	840
QY	841	CTCACAAATATGTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC	900
DB	841	CTCACAAATATGTAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAGAGTCTACCTCC	900
QY	901	CTATCCAGCATCCCGGACTCTTCCCAATTAAGGACCCCTTCAACCCAAATGG	960
DB	901	CTATCCAGCATCCCGGACTCTTCCCAATTAAGGACCCCTTCAACCCAAATGG	960
QY	961	TCCAAAGGAGTAGACAAAGGGAACAGTGAACCAAGAGTGCCTAATTTCCCAAT	1020
DB	961	TCCAAAGGAGTAGACAAAGGGAACAGTGAACCAAGAGTGCCTAATTTCCCAAT	1020
QY	1021	TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTTGGCCACGACAGTGCATGTGCT	1080

DB	1021	TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGCCCCAGCCAGAGTGCATGTGCT	1080
QY	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAATCCT	1140
DB	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAATCCT	1140
QY	1141	GATGCTATATTTGRTGTTTTTACAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT	1200
DB	1141	GATGCTATATTTGRTGTTTTTACAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT	1200
QY	1201	ATATATGTCATGCTGCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCACCAATATCTGC	1260
DB	1201	ATATATGTCATGCTGCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCACCAATATCTGC	1260
QY	1261	AGCTTGAGGTTTGGCGATCTCTGTATCTCAGTCAGTCAATGAGTANGGATGACAAACA	1320
DB	1261	AGCTTGAGGTTTGGCGATCTCTGTATCTCAGTCAGTCAATGAGTANGGATGACAAACA	1320
QY	1321	GAGGAAAGANAATGATTTCCCAAGGCCAGCAGCAGTTCCTCAGTCTASACCTCATTTG	1380
DB	1321	GAGGAAAGANAATGATTTCCCAAGGCCAGCAGCAGTTCCTCAGTCTASACCTCATTTG	1380
QY	1381	GGGACACAGAAATCAGTAACATGGGAGATTTGCTGTCAGACATTTTCTAACTTGTGTC	1440
DB	1381	GGGACACAGAAATCAGTAACATGGGAGATTTGCTGTCAGACATTTTCTAACTTGTGTC	1440
QY	1441	TASAGGACTTAAGGAAACTASAGAAAGAAATCTATVGAATTAATCTCAATGATGTCCACATA	1500
DB	1441	TASAGGACTTAAGGAAACTASAGAAAGAAATCTATVGAATTAATCTCAATGATGTCCACATA	1500
QY	1501	ACACAGGGGAGGGAAGAAAATCTCTATCCTCTTTCTGGAGAGACTAAGGAGGCAATGAG	1560
DB	1501	ACACAGGGGAGGGAAGAAAATCTCTATCCTCTTTCTGGAGAGACTAAGGAGGCAATGAG	1560
QY	1561	GAGCGTGCCTCTCTGTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCGTAAGTT	1620
DB	1561	GAGCGTGCCTCTCTGTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCGTAAGTT	1620
QY	1621	TATCACTCAGTGCAGTGCAGACATTAAGAAAACCTTCAAAAGTCTCGCTAGGCCCGGAG	1680
DB	1621	TATCACTCAGTGCAGTGCAGACATTAAGAAAACCTTCAAAAGTCTCGCTAGGCCCGGAG	1680
QY	1681	CAAAAATTAGAAACCTTATTTGAACTTTGGAACCTTTGGCAACTCGCTTATTAAGATCAGGAG	1740
DB	1681	CAAAAATTAGAAACCTTATTTGAACTTTGGAACCTTTGGCAACTCGCTTATTAAGATCAGGAG	1740
QY	1741	GAGCAGCGGACAGGACAAACGGGATTAAGAAAAGGCAACCGCTTTAGTCAATGACCT	1800
DB	1741	GAGCAGCGGACAGGACAAACGGGATTAAGAAAAGGCAACCGCTTTAGTCAATGACCT	1800
QY	1801	CAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGGAAGAAAGCTGGGCAAAATGAATGCTAA	1860
DB	1801	CAGGCAAGTGGACTTTTGGAGGCTCTGAAAAGGGAAGAAAGCTGGGCAAAATGAATGCTAA	1860
QY	1861	TAGGGCTTCTTCAGTGCCTCTACAAGACACTTTTAAAGAGATTTGTCCAAGTAGAAG	1920
DB	1861	TAGGGCTTCTTCAGTGCCTCTACAAGACACTTTTAAAGAGATTTGTCCAAGTAGAAG	1920
QY	1921	TAAAGCCCTCTCTGCTCAATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCC	1980
DB	1921	TAAAGCCCTCTCTGCTCAATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCC	1980
QY	1981	CAGGGACAAAGGTCTTTTGAAGTCAAGACCACTAACACAGATGATCAGCAGCAGGACTG	2040
DB	1981	CAGGGACAAAGGTCTTTTGAAGTCAAGACCACTAACACAGATGATCAGCAGCAGGACTG	2040
QY	2041	AGGGTCTCTGGGCAAGCGCCATCCATGCTCAACCTCAAGAGCCCTGGGTATGCTT	2100
DB	2041	AGGGTCTCTGGGCAAGCGCCATCCATGCTCAACCTCAAGAGCCCTGGGTATGCTT	2100
QY	2101	GACCAATGAGGCGCAGGAGGTTGCTCTCTGGACACTGGTGGCTTCTTCTAGTCTTACT	2160
DB	2101	GACCAATGAGGCGCAGGAGGTTGCTCTCTGGACACTGGTGGCTTCTTCTAGTCTTACT	2160

QY 2161 CTTCTGTCCGGCAACCTGTCCTCCAGATCTGTCACTATTCTCGAGGGGTCCNTAAGACG 2220  
DB 2161 CTTCTGTCCGGCAACCTGTCCTCCAGATCTGTCACTATTCTCGAGGGGTCCNTAAGACG 2220  
QY 2221 GGCAGTCACTAGATACATTTTTCAGCCACCTAAAGTATGAATCGGGAGCTTTATCTTT 2280  
DB 2221 GGCAGTCACTAGATACATTTTTCAGCCACCTAAAGTATGAATCGGGAGCTTTATCTTT 2280  
QY 2281 TCACATGCTTTTCTCAATATATGCTTTGAAAGCCCACTACCTTTGTTGAGGAGAGACATCTTA 2340  
DB 2281 TCACATGCTTTTCTCAATATATGCTTTGAAAGCCCACTACCTTTGTTGAGGAGAGACATCTTA 2340  
QY 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACAACCCGTTTGTGTGCC 2400  
DB 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACAACCCGTTTGTGTGCC 2400  
QY 2401 CTTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAATATGGACGAGC 2460  
DB 2401 CTTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAATATGGACGAGC 2460  
QY 2461 CAAAGAAATGCCCGTCTGTTCAAGTTAAACTAAAGGATTCACCTTCCCTTCCCTACCAA 2520  
DB 2461 CAAAGAAATGCCCGTCTGTTCAAGTTAAACTAAAGGATTCACCTTCCCTTCCCTACCAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTAGTAAACCATGATTAACCTCCCTGCAAGTAACTCCGTTAGTGGATTGAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCATGATTAACCTCCCTGCAAGTAACTCCGTTAGTGGATTGAG 2640  
QY 2641 GAGGCACAGAACCCAGTGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAACAATGG 2700  
DB 2641 GAGGCACAGAACCCAGTGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAACAATGG 2700  
QY 2701 AGGCCGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCTCCAAATAC 2760  
DB 2701 AGGCCGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCTCCAAATAC 2760  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTGGACCTTMAAGATCCCTTCTGTCATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGGTTTACASTCTGGACCTTMAAGATCCCTTCTGTCATCCCTG 2820  
QY 2821 TACATCTGACTCTCAATCTTGTGCTTTGCAAGATACCTTCAAAACCCARCATCTCAAC 2880  
DB 2821 TACATCTGACTCTCAATCTTGTGCTTTGCAAGATACCTTCAAAACCCARCATCTCAAC 2880  
QY 2881 TCACCTGGACTRTTTTATACCCCAAGGTTTCAGGGATAGYCCCCATCTATTGTCAGGAGCAT 2940  
DB 2881 TCACCTGGACTRTTTTATACCCCAAGGTTTCAGGGATAGYCCCCATCTATTGTCAGGAGCAT 2940  
QY 2941 TAGCCCAAGACTGAGYCARVTMTCACTCTGGACACTCTTGTCTTCRGTAKGTGGATG 3000  
DB 2941 TAGCCCAAGACTGAGYCARVTMTCACTCTGGACACTCTTGTCTTCRGTAKGTGGATG 3000  
QY 3001 ATTTACTTTTTRGTCYCCVTTTCAGAAACCTTGTGTCATCAAGCCACCAAGCRCTCTTMA 3060  
DB 3001 ATTTACTTTTTRGTCYCCVTTTCAGAAACCTTGTGTCATCAAGCCACCAAGCRCTCTTMA 3060  
QY 3061 ATTTCTCGYACCTGTGCTACAGGTTTCCAAACARSARGCTCARTCTGCTCACAGC 3120  
DB 3061 ATTTCTCGYACCTGTGCTACAGGTTTCCAAACARSARGCTCARTCTGCTCACAGC 3120  
QY 3121 AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
DB 3121 AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGCACCAAGGCCCTCAGTGAGGAAYRYA 3180  
QY 3181 TCCAGCCATATCTGCTTATCTCATCTCAATCYCAAAACCTTAAAGCAATTAAGRRRTTCTTTG 3240  
DB 3181 TCCAGCCATATCTGCTTATCTCATCTCAATCYCAAAACCTTAAAGCAATTAAGRRRTTCTTTG 3240

QY 3241 GCRTAAYAGGYTTCTGCCGAANWATGGATTCCCCAGGTWTGGCBAAATAGCCAGGYCATTA 3300  
DB 3241 GCRTAAYAGGYTTCTGCCGAANWATGGATTCCCCAGGTWTGGCBAAATAGCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGGAACACTCAGAAAGCAATACCCTATTARTAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAACACTCAGAAAGCAATACCCTATTARTAGATGGAYAMCTGAA 3360  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYT 3420  
DB 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYT 3420  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAVRTCACAGAAAACACAGRAAYAGCTCTRGGA 3480  
DB 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAVRTCACAGAAAACACAGRAAYAGCTCTRGGA 3480  
QY 3481 GTTCCTTACACAGRTCCRAGGGAYGAGCTTGCACCCYRTGGCRYACCTGASTAAGGAAYT 3540  
DB 3481 GTTCCTTACACAGRTCCRAGGGAYGAGCTTGCACCCYRTGGCRYACCTGASTAAGGAAYT 3540  
QY 3541 GATGTAGTGGCAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTACGATYKTA 3600  
DB 3541 GATGTAGTGGCAAGGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTACGATYKTA 3600  
QY 3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB 3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
QY 3661 GTGAAYRGCATACCTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGAACVGTGTTACTTAA 3720  
DB 3661 GTGAAYRGCATACCTCACTGTCTAAAGGAGACTTGTGGCTGTGACAGAACVGTGTTACTTAA 3720  
QY 3721 TRTCAGGCTCTATTACTTGAAGGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
DB 3721 TRTCAGGCTCTATTACTTGAAGGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
QY 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATACTGTCTCAACAAATTAAT 3840  
DB 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATACTGTCTCAACAAATTAAT 3840  
QY 3841 TCTCAAACTATGCCACTTCGAGGGGACCTTGTAGAGTTTCVTTGACTGATCCYACACTT 3900  
DB 3841 TCTCAAACTATGCCACTTCGAGGGGACCTTGTAGAGTTTCVTTGACTGATCCYACACTT 3900  
QY 3901 CAACCTGTATACCTGATGAGGAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
DB 3901 CAACCTGTATACCTGATGAGGAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
QY 3961 AGTGTGATGATTAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
DB 3961 AGTGTGATGATTAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
QY 4021 GCTRGCAAGAACTAATAGCCYTCVTTGKGGCACTAGAAATTAGGAGAGRAAAAAGGGYAAA 4080  
DB 4021 GCTRGCAAGAACTAATAGCCYTCVTTGKGGCACTAGAAATTAGGAGAGRAAAAAGGGYAAA 4080  
QY 4081 TATATATACAGACTCTTARTATGCTYACCTAGTCNTCCATGCCCATATGATGSR 4140  
DB 4081 TATATATACAGACTCTTARTATGCTYACCTAGTCNTCCATGCCCATATGATGSR 4140  
QY 4141 AGAAGGGGAATTCCTTAACCTTCYAGRGAAACCTTATCAMACATCAGGAAGCATTAGGAR 4200  
DB 4141 AGAAGGGGAATTCCTTAACCTTCYAGRGAAACCTTATCAMACATCAGGAAGCATTAGGAR 4200  
QY 4201 ATTATTAATGTCGCTACAGAAACCTTAGAGGTTGMAAGTCTTACACTCYGGGGTCATCA 4260  
DB 4201 ATTATTAATGTCGCTACAGAAACCTTAGAGGTTGMAAGTCTTACACTCYGGGGTCATCA 4260  
QY 4261 NAAAGGAAAGRAAGGAAATASAGRGAAATGCAAGACATATTTGAAGCMAAAGAGC 4320  
DB 4261 NAAAGGAAAGRAAGGAAATASAGRGAAATGCAAGACATATTTGAAGCMAAAGAGC 4320  
QY 4321 TGCAGGCAAGGACCTCCATCTAGNAATGCTTATTAATAAATCCCTTCCCTTAGTATAGGGTAATCC 4380

Db 4321 TGCAAGCAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTAGGATATCC 4380  
Qy  
Db 4381 CTTCCGGGAACCAAGCCCAAGTACTCAGCAGAGGAAACAGAAATGGGGACCTCAGAGG 4440  
Qy  
Db 4381 CTTCCGGGAACCAAGCCCAAGTACTCAGCAGAGGAAACAGAAATGGGGACCTCAGAGG 4440  
Qy 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGAAATACTTTGCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGAAATACTTTGCTGCAAC 4500  
Qy 4501 TATCCAAATGAAATTAATTAACCCCTTCAATCACTTTCACTTAGGCAATCAGTAGCAC 4560  
Db 4501 TATCCAAATGAAATTAATTAACCCCTTCAATCACTTTCACTTAGGCAATCAGTAGCAC 4560  
Qy 4561 CCATCARATGCGCAAAATCAATTTACTGACACAGGCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGCGCAAAATCAATTTACTGACACAGGCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGGCTGTGAATGTGCCARAAATAATCCCTGCTCTATGCGCAAGCTCCTTC 4680  
Db 4621 AKTCAGGGCTGTGAATGTGCCARAAATAATCCCTGCTCTATGCGCAAGCTCCTTC 4680  
Qy 4681 AGGARAACAARAACAGGCCATTAACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db 4681 AGGARAACAARAACAGGCCATTAACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Qy 4741 CCCAAACCTCAGGGATTTAGTACTACTAGTCTGGGTARATCTTTCAAGGTTGGCA 4800  
Db 4741 CCCAAACCTCAGGGATTTAGTACTACTAGTCTGGGTARATCTTTCAAGGTTGGCA 4800  
Qy 4801 RAGGCCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAA 4860  
Db 4801 RAGGCCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAA 4860  
Qy 4861 ATAATTCAGATTCGAGTTCCCGAGGCTTACAGAGTGACATAGACCCCTGCTTCCAG 4920  
Db 4861 ATAATTCAGATTCGAGTTCCCGAGGCTTACAGAGTGACATAGACCCCTGCTTCCAG 4920  
Qy 4921 GCCACGTACCCAGGAGTATCCAGGCTTAGGTATAGATACATCACTACCTGCGCC 4980  
Db 4921 GCCACGTACCCAGGAGTATCCAGGCTTAGGTATAGATACATCACTACCTGCGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCAGGGAAAGTCGAGAAAATGAATGAAAYACTCAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGGAAAGTCGAGAAAATGAATGAAAYACTCAAGGACATCTA 5040  
Qy 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGTCCTATAGCCCTTAAA 5100  
Db 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGTCCTATAGCCCTTAAA 5100  
Qy 5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTAGGAAG 5160  
Db 5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTAGGAAG 5160  
Qy 5161 CCCTTCATAACCAATGACCTTGTCTTGACCCCAAGACGCGCAACTTAGTTGCGACATCA 5220  
Db 5161 CCCTTCATAACCAATGACCTTGTCTTGACCCCAAGACGCGCAACTTAGTTGCGACATCA 5220  
Qy 5221 CTTCTTAGGCCAAATATCAACAGTTCTTAAAACATTAACAGGAAACCTATCCCTGAGAAG 5280  
Db 5221 CTTCTTAGGCCAAATATCAACAGTTCTTAAAACATTAACAGGAAACCTATCCCTGAGAAG 5280  
Qy 5281 AGGAAAAGAACATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340  
Db 5281 AGGAAAAGAACATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATYACCCCAACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATYACCCCAACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTGAGTCTTGGATACATCACTGAGTCAAAATCCTGGATAC 5460

Db 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACTGAAAAATCCAGGAGACAAACCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACTGAAAAATCCAGGAGACAAACCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
Qy 5521 ATTGTGCGCTGCTCTTTCAAAACAAACACGAGGAGAAAGTAATATAAATCATATAATCCCC 5580  
Db 5521 ATTGTGCGCTGCTCTTTCAAAACAAACACGAGGAGAAAGTAATATAAATCATATAATCCCC 5580  
Qy 5581 ATGGSCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTTCACTCTCACT 5640  
Db 5581 ATGGSCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGGAGA 5700  
Qy 5701 ATGACGGTCCCGGAAATATTGATGCCCATGATAGAGTCTTTSTAAGGAAACCCCC 5760  
Db 5701 ATGACGGTCCCGGAAATATTGATGCCCATGATAGAGTCTTTSTAAGGAAACCCCC 5760  
Qy 5761 ACCTTCACCTGCCACACCCCATATGCGCGCTGATGACAGTAGTCCCTTTCGATG 5820  
Db 5761 ACCTTCACCTGCCACACCCCATATGCGCGCTGATGACAGTAGTCTTTTCGATG 5820  
Qy 5821 CATGCAAAATACATCATTTATTGGACAGGAAAAATGATTAACTCTAGTTGCTCTGAGGACTT 5880  
Db 5821 CATGCAAAATACATCATTTATTGGACAGGAAAAATGATTAACTCTAGTTGCTCTGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTTACCCAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTTACCCAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAACATGTAAAGAGTAACTCTCCCAACTCAACCCGCTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAACATGTAAAGAGTAACTCTCCCAACTCAACCCGCTACAT 6000  
Qy 6001 GGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTAACC 6060  
Db 6001 GGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTAACC 6060  
Qy 6061 CATATCTGCGCTGTGAGCTTATTAATAACAACCTCTGCGGCTCATGAGGCTCGGCC 6120  
Db 6061 CATATCTGCGCTGTGAGCTTATTAATAACAACCTCTGCGGCTCATGAGGCTCGGCC 6120  
Qy 6121 CAACCCCTACTAACTGTGGATATGCTCCCTCGAACTTCARGCCATATGTTCAATC 6180  
Db 6121 CAACCCCTACTAACTGTGGATATGCTCCCTCGAACTTCARGCCATATGTTCAATC 6180  
Qy 6181 CCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACCACTTCGCTTTTAGTA 6240  
Db 6181 CCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACCACTTCGCTTTTAGTA 6240  
Qy 6241 GGAACCTCTGTTTCCAAATSTGGAAATAAACCACTTCAAACCTCACTGTGTAAATTTT 6300  
Db 6241 GGAACCTCTGTTTCCAAATSTGGAAATAAACCACTTCAAACCTCACTGTGTAAATTTT 6300  
Qy 6301 AGCAATACTACATACACACCACTCCCATGATCAGGTGGGTAACTCCCTCCACACAA 6360  
Db 6301 AGCAATACTACATACACACCACTCCCATGATCAGGTGGGTAACTCCCTCCACACAA 6360  
Qy 6361 ATAGTCTGCGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
Db 6361 ATAGTCTGCGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTG 6420  
Qy 6421 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGTGCCCCCYATGRCATCTAC 6480  
Db 6421 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTAGTGTGCCCCCYATGRCATCTAC 6480  
Qy 6481 ACTGAAACAGATTTTATACAGTTATGTCATATTAAGCCCCCGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGAAACAGATTTTATACAGTTATGTCATATTAAGCCCCCGCAACAAAGAGTACCCATT 6540



QY 6541 CTTCTTTTGTATTAGGAGCAGAGTGTAGTGTCACTAGTACTGGCATTTGGCGTATC 6600  
Db 6541 CTTCTTTTGTATTAGGAGCAGAGTGTAGTGTCACTAGTACTGGCATTTGGCGTATC 6600  
QY 6601 ACNACCTCTACTAGTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Db 6601 ACNACCTCTACTAGTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
QY 6661 GTCGCCGACTCCCTGTGTCACTTGCAGAGTCACTTAACTCCCTAGCAGCATGTCCTT 6720  
Db 6661 GTCGCCGACTCCCTGTGTCACTTGCAGAGTCACTTAACTCCCTAGCAGCATGTCCTT 6720  
QY 6721 CRAAATCGAAGAGCTTTAGACTYGTAAACCGCTGAGAGGGGGAACCTGTTATTATTTTA 6780  
Db 6721 CRAAATCGAAGAGCTTTAGACTYGTAAACCGCTGAGAGGGGGAACCTGTTATTATTTTA 6780  
QY 6781 GGGGAAGATGCTGTTATTATGTTATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Db 6781 GGGGAAGATGCTGTTATTATGTTATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
QY 6841 ATTCSAGATCGAATACAACTAGCAGAGAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
Db 6841 ATTCSAGATCGAATACAACTAGCAGAGAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
QY 6901 CTCAGCCATGATGCTGCTGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTCAGCCATGATGCTGCTGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
QY 6961 CTACTCTCTTTGGACCCCTGATCTTACCTCTTGTAACTGTTTAACTGTTTCCAGAATC 7020  
Db 6961 CTACTCTCTTTGGACCCCTGATCTTACCTCTTGTAACTGTTTAACTGTTTCCAGAATC 7020  
QY 7021 GAAGCTGTAACTCAAACTAGGAGGAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 7080  
Db 7021 GAAGCTGTAACTCAAACTAGGAGGAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 7080  
QY 7081 CCCCTGGACCGGCTGTAGTACCCAGCATCTGATGTTATGATCAATCAAGGACCCCTCTCT 7140  
Db 7081 CCCCTGGACCGGCTGTAGTACCCAGCATCTGATGTTATGATCAATCAAGGACCCCTCTCT 7140  
QY 7141 GAGGAATCTCAGCTGCAAACTCTACTACGCCCAATTCAGCAGGAGCACTAGTACGAGC 7200  
Db 7141 GAGGAATCTCAGCTGCAAACTCTACTACGCCCAATTCAGCAGGAGCACTAGTACGAGC 7200  
QY 7201 GGTGTGCGGCAACCTCCCAACAGCACTTACTAGTCTTCTGTTGAGATGGGAGCTGAGA 7260  
Db 7201 GGTGTGCGGCAACCTCCCAACAGCACTTACTAGTCTTCTGTTGAGATGGGAGCTGAGA 7260  
QY 7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCTTAGTGGGAGGTG 7320  
Db 7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCTTAGTGGGAGGTG 7320  
QY 7321 ACCAATCCACTTTAAACAGGGGCTTGCACTTACTAGTCTTCTGTTGAGATGGGAGCTGAGA 7380  
Db 7321 ACCAATCCACTTTAAACAGGGGCTTGCACTTACTAGTCTTCTGTTGAGATGGGAGCTGAGA 7380  
QY 7381 CTCCTAAATGCTAAATAGGACAGAGGAGTAAAGAACTAGCAATCATYTTATTGTC 7440  
Db 7381 CTCCTAAATGCTAAATAGGACAGAGGAGTAAAGAACTAGCAATCATYTTATTGTC 7440  
QY 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTTCAGGCGGCA 7500  
Db 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTTCAGGCGGCA 7500  
QY 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
QY 7561 TCTATTAAATCTTGACCTGCR 7582  
Db 7561 TCTATTAAATCTTGACCTGCR 7582

RESULT 2  
BD267487LOCUS DEFINITION 7582 bp DNA linear PAT 17-JUL-2003  
Endogenous nucleic fragment associated with an autoimmune disease,  
marking method and reagent.ACCESSION BD267487  
VERSION 1 GI:33077255  
KEYWORDS JP 2002534980-A/28.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 7582)  
AUTHORS Baccalini, G., Mallet, F. and Voisset, C.TITLE Endogenous nucleic fragment associated with an autoimmune disease,  
marking method and reagentJOURNAL Patient: JP 2002534980-A 28 22-OCT-2002;  
BIO MERIEUXCOMMENT OS Homo sapiens (human)  
DN JP 2002534980-A/28PP 22-OCT-2002  
PI 21-JAN-1999 FR 99/00888PR GLAUCIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC  
C12N15/09, C07K14/15, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, PCG01N33/564,  
PC G01N33/566, G01N33/569, G01N33/58//C12P21/02, C12N15/00 CCEndogenous nucleic fragment associated with an autoimmune CC  
disease,

CC marking method and reagent

FH Key source

FT Location/Qualifiers  
1. 7582

/organism="Homo sapiens (human)"

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 98.8%; Score 7489.6; DB 6; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGTCC 60

Db 1 CAACAATCGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGTCC 60

QY 61 CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCATCTTAAATCTTGCAACTGCA 120

Db 61 CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCATCTTAAATCTTGCAACTGCA 120

QY 121 CTCCTTCTGCTCCATGTTTCTTACCGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCATCTGC 180

Db 121 CTCCTTCTGCTCCATGTTTCTTACCGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCATCTGC 180

QY 181 TGTGTTGCCACACCGCAGACCTGCGCTGATCTCCATCCCTCTGATCTCGAGGCTGTC 240

Db 181 TGTGTTGCCACACCGCAGACCTGCGCTGATCTCCATCCCTCTGATCTCGAGGCTGTC 240

QY 241 CGCTGTCTCTCTGATCCAGGAGCGGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTGC 300

Db 241 CGCTGTCTCTCTGATCCAGGAGCGGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTGC 300

QY 301 CATTGTGNCCTGACGGCTAAAGTCTGGGTTTGTCTTAAATGAGCTGAACACTANTCACT 360

Db 301 CATTGTGNCCTGACGGCTAAAGTCTGGGTTTGTCTTAAATGAGCTGAACACTANTCACT 360

QY 361 GGGTTCATGCTTCTTCTGTGACCCACCGGCTTCTAATAKAACCTATAACACTTACCACA 420

Db 361 GGGTTCATGCTTCTTCTGTGACCCACCGGCTTCTAATAKAACCTATAACACTTACCACA 420

421 TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAACGAATCTCCAGGTGAGAGATAC 480  
421 TGGCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAACGAATCTCCAGGTGAGAGATAC 480  
481 GARGCTTGCACCATCTTGGAAAGCGGCTGCTACCRCTTGGAAAGTGGTTTCAACCATTC 540  
481 GARGCTTGCACCATCTTGGAAAGCGGCTGCTACCRCTTGGAAAGTGGTTTCAACCATTC 540  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCBACCAASRAAGGACATCC 600  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCBACCAASRAAGGACATCC 600  
601 MAAGTGATGSGAAACGTTCCCGCAAGACAAACGCGCTTAAGAGCTATTTCTGGARAAT 660  
601 MAAGTGATGSGAAACGTTCCCGCAAGACAAACGCGCTTAAGAGCTATTTCTGGARAAT 660  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGAGTGCC 720  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGAGTGCC 720  
721 GCTTGGCACTCTGAGGGAAGTATAAATTATAACACATCTTACAGCTAGACTCTTTTG 780  
721 GCTTGGCACTCTGAGGGAAGTATAAATTATAACACATCTTACAGCTAGACTCTTTTG 780  
781 TAGAAAGGCAAAATGAGTGAAGTGCCATAGTACAAACCTTTCTTTTCAATTAAAGACAA 840  
781 TAGAAAGGCAAAATGAGTGAAGTGCCATAGTACAAACCTTTCTTTTCAATTAAAGACAA 840  
841 CTCAAAATATGTAAGAAAGTGAATTTATGCCCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
841 CTCAAAATATGTAAGAAAGTGAATTTATGCCCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
901 CTATCCAGCATCCCGACTCTTCCCGCAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 960  
901 CTATCCAGCATCCCGACTCTTCCCGCAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 960  
961 TCCAAAGGAGATAGACAAAGGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020  
961 TCCAAAGGAGATAGACAAAGGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTTGGCCCGCAGCAGAGTGCATGTCYT 1080  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTTGGCCCGCAGCAGAGTGCATGTCYT 1080  
1081 TTTTCTCCAGCATCCCGACTCTTCCCGCAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140  
1081 TTTTCTCCAGCATCCCGACTCTTCCCGCAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140  
1141 GATGGCTATATGTTGTTTACAGGCTTGAAGCAATTTCTTGTATCTGACATGAGAGAT 1200  
1141 GATGGCTATATGTTGTTTACAGGCTTGAAGCAATTTCTTGTATCTGACATGAGAGAT 1200  
1201 ATATATGCTCACTGCTAAATCAGACACTAACCCCAATGAGAAAGTGCACCACTGTC 1260  
1201 ATATATGCTCACTGCTAAATCAGACACTAACCCCAATGAGAAAGTGCACCACTGTC 1260  
1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320  
1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320  
1321 GAGGAAGANAATGATTTCCCAAGGCTGAGGAGTTCAGGCTTASACCCCTCATTTG 1380  
1321 GAGGAAGANAATGATTTCCCAAGGCTGAGGAGTTCAGGCTTASACCCCTCATTTG 1380  
1381 GGGACACAGAAATCAGTAAACATGAGGAGTGGTCTGCAGACATTTGCTAACTTGTGTGC 1440  
1381 GGGACACAGAAATCAGTAAACATGAGGAGTGGTCTGCAGACATTTGCTAACTTGTGTGC 1440  
1441 TASAAGGACTAAGGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCATTA 1500  
1441 TASAAGGACTAAGGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCATTA 1500  
1501 ACACAGGGAAGGGAAGAAATCTTACTGCGCTTTCTGGAGAGACTAAGGGAGGCAATTGAG 1560

1501 ACACAGGGAAGGGAAGAAATCTTACTGCGCTTTCTGGAGAGACTAAGGGAGGCAATTGAG 1560  
1561 GAGGGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAACTCTTAAAGCGTAAGTT 1620  
1561 GAGGGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAACTCTTAAAGCGTAAGTT 1620  
1621 TATCACTCAGTCTGAGTGCAGACATTTAGAAAAAACTTTCAAAAGTCTGCGTAGGCCGGAG 1680  
1621 TATCACTCAGTCTGAGTGCAGACATTTAGAAAAAACTTTCAAAAGTCTGCGTAGGCCGGAG 1680  
1681 CAAAACTTTAGAAAACTTGAACCTTGGCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
1681 CAAAACTTTAGAAAACTTGAACCTTGGCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
1741 GAGCAGGCGGACACGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCAATGACCTT 1800  
1741 GAGCAGGCGGACACGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCAATGACCTT 1800  
1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
1861 TAGGGCTTGTTCAGTGGGCTTCAAGGACACTTTAAAAAAGGATTTGTCCAAGTAGAAG 1920  
1861 TAGGGCTTGTTCAGTGGGCTTCAAGGACACTTTAAAAAAGGATTTGTCCAAGTAGAAG 1920  
1921 TAAGCGGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCCC 1980  
1921 TAAGCGGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCCC 1980  
1981 CAGGGAACAAAGGCTCTTTTGAAGTCAAGAACCTAAACAGATGATCCAGCAGCAGACTG 2040  
1981 CAGGGAACAAAGGCTCTTTTGAAGTCAAGAACCTAAACAGATGATCCAGCAGCAGACTG 2040  
2041 AGGGTGCCTGGGGAAGGCCATCCCATGCCATCACCTTCAAGAGCCCTTGGGTATGCTT 2100  
2041 AGGGTGCCTGGGGAAGGCCATCCCATGCCATCACCTTCAAGAGCCCTTGGGTATGCTT 2100  
2101 GACCATGAGGCGCAGGAAGGTTGCTCTGGACACTGGTGGGCTTCTTACTCTTACT 2160  
2101 GACCATGAGGCGCAGGAAGGTTGCTCTGGACACTGGTGGGCTTCTTACTCTTACT 2160  
2161 CTCTCTCTCCCGCAACTGTCTCTCCAGATCTGTCACTATTTCTGAGGGGCTCCTAAGAGC 2220  
2161 CTCTCTCTCCCGCAACTGTCTCTCCAGATCTGTCACTATTTCTGAGGGGCTCCTAAGAGC 2220  
2221 GGCACTCAGTATAGATCTTTTCCAGGCACTAAGTTTATGNACTGGGAGGCTTTTCTTTT 2280  
2221 GGCACTCAGTATAGATCTTTTCCAGGCACTAAGTTTATGNACTGGGAGGCTTTTCTTTT 2280  
2281 TCACATGCTTTTCTAATTTATGTTAAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
2281 TCACATGCTTTTCTAATTTATGTTAAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
2341 GCAAAAGCAGGCGCATTATACACCTGAAATAGGAGAGAAACACCCGTTTGTGTGTC 2400  
2341 GCAAAAGCAGGCGCATTATACACCTGAAATAGGAGAGAAACACCCGTTTGTGTGTC 2400  
2401 CTTGCTTGGAGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGAATATGAGCAGC 2460  
2401 CTTGCTTGGAGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGAATATGAGCAGC 2460  
2461 CAAAGATGCGGCTCTGTTTCAAGTTTAACTAAAGGATTCACCTTCTTCTTCTTCTTCTTCT 2520  
2461 CAAAGATGCGGCTCTGTTTCAAGTTTAACTAAAGGATTCACCTTCTTCTTCTTCTTCTTCT 2520  
2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTCCTTCTTCTTCTTCTTCTTCTTCT 2580  
2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTCCTTCTTCTTCTTCTTCTTCTTCT 2580  
2581 AAGCCCAAGGCTTAGTAAACCAATGATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2640

Db	2581	AAGCCCAAGGCTTAGTAAACCAACATGATTAACCTCCCTGCGATTAATCCCTAGTGGATTGAG	2640
Qy	2641	GAGGCACAGAAACCCAGTGGGAGCAGTGGAGGGTTAGTGCAGATCTCAGGATTAACAATGG	2700
Db	2641	GAGGCACAGAAACCCAGTGGGAGCAGTGGAGGGTTAGTGCAGATCTCAGGATTAACAATGG	2700
Qy	2701	AGGCGGTGTCTCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCCAAAATAC	2760
Db	2701	AGGCGGTGTCTCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCCAAAATAC	2760
Qy	2761	CAGAGGACAGAGTGGTTTACSTCCCTGGACCTTMAGGATGCTCTTCTGTGATCCCTG	2820
Db	2761	CAGAGGACAGAGTGGTTTACSTCCCTGGACCTTMAGGATGCTCTTCTGTGATCCCTG	2820
Qy	2821	TACATCCTGACTCTCAATCTCTTGTGTGCTTTGAGATACCTTCAAAACCCARCATCTCAAC	2880
Db	2821	TACATCCTGACTCTCAATCTCTTGTGTGCTTTGAGATACCTTCAAAACCCARCATCTCAAC	2880
Qy	2881	TCACCTGGACTRRTTTTACCCCAAGGTTTACGGGATAGYCCCATCTATTTGGCCAGGCAT	2940
Db	2881	TCACCTGGACTRRTTTTACCCCAAGGTTTACGGGATAGYCCCATCTATTTGGCCAGGCAT	2940
Qy	2941	TAGCCCAAGACTTGCAGYCARITYMTACCTGGACACTCTTGTCTCTCRGTAKGTGGATG	3000
Db	2941	TAGCCCAAGACTTGCAGYCARITYMTACCTGGACACTCTTGTCTCTCRGTAKGTGGATG	3000
Qy	3001	ATTTACTTTTTRGCGCCVRRTTTCAAGAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA	3060
Db	3001	ATTTACTTTTTRGCGCCVRRTTTCAAGAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA	3060
Qy	3061	ATTTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACCSARAGCTCARCTCTGTCTCACAGC	3120
Db	3061	ATTTCTCTCGYACCTGTGGCTTACAWGGTTTCCAAACCSARAGCTCARCTCTGTCTCACAGC	3120
Qy	3121	AGSTTAAATACTTAGRCHTAARATTTCCAAAGGCAACARGGCCCTCAGTGAGGAAVRYA	3180
Db	3121	AGSTTAAATACTTAGRCHTAARATTTCCAAAGGCAACARGGCCCTCAGTGAGGAAVRYA	3180
Qy	3181	TCCAGCCTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCACTAAGGRRTTCTTGG	3240
Db	3181	TCCAGCCTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCACTAAGGRRTTCTTGG	3240
Qy	3241	GCRTAAYAGGYTTCTGCGGAATGATTCCTCAGAGTWTGGCRAAATAGCCAGGYCATTA	3300
Db	3241	GCRTAAYAGGYTTCTGCGGAATGATTCCTCAGAGTWTGGCRAAATAGCCAGGYCATTA	3300
Qy	3301	WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTATTAAGATGAGVAMCTGAA	3360
Db	3301	WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTATTAAGATGAGVAMCTGAA	3360
Qy	3361	GYMRAAGTGGCTTTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCCAAGTCTTAAGYT	3420
Db	3361	GYMRAAGTGGCTTTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCCAAGTCTTAAGYT	3420
Qy	3421	TGCCAACRGGGCAAGACTTTTSTTYATVATTCACAGAAACCAACAGRAAYAGCTCTRGGG	3480
Db	3421	TGCCAACRGGGCAAGACTTTTSTTYATVATTCACAGAAACCAACAGRAAYAGCTCTRGGG	3480
Qy	3481	GTCTCTTACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRVACCTGASTAAGGAAAYT	3540
Db	3481	GTCTCTTACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRVACCTGASTAAGGAAAYT	3540
Qy	3541	GATGTAGTGGCAAGGGTTGRCVTCATTTGTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA	3600
Db	3541	GATGTAGTGGCAAGGGTTGRCVTCATTTGTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA	3600
Qy	3601	GTATCTGAAGCAGTTTAAATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Db	3601	GTATCTGAAGCAGTTTAAATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Qy	3661	GTGAAYRGCACTACTCTGCTAAGGAGACTTGTGGCTGTGCAGAACYGTTTTACTTTAA	3720
Db	3661	GTGAAYRGCACTACTCTGCTAAGGAGACTTGTGGCTGTGCAGAACYGTTTTACTTTAA	3720
Qy	3721	TRTCAGGCTCTATTTACTTGAAGGCCAGTGTGTCRACCTGTGCACTTGTGCAACTCTTAAAC	3780
Db	3721	TRTCAGGCTCTATTTACTTGAAGGCCAGTGTGTCRACCTGTGCACTTGTGCAACTCTTAAAC	3780
Qy	3781	CCAGYCNCAATTTCTTCCAGACCAATGAAGAAAGATARAAYATACTGTCAACAARTATTT	3840
Db	3781	CCAGYCNCAATTTCTTCCAGACCAATGAAGAAAGATARAAYATACTGTCAACAARTATTT	3840
Qy	3841	TCTCAAAACCTATGCACTCGAGGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYACCTT	3900
Db	3841	TCTCAAAACCTATGCACTCGAGGGGACCTTCTTGTAGAGTTCCTTGTGACTGATCCYACCTT	3900
Qy	3901	CAACTTGTATACATGATGGAAGTTCCTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGC	3960
Db	3901	CAACTTGTATACATGATGGAAGTTCCTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGC	3960
Qy	3961	AGTGGTCAGTGATTAATGGAATATTTGAAAGTAATCCCTCCTCACCTCCAGGAACTAGTCTYA	4020
Db	3961	AGTGGTCAGTGATTAATGGAATATTTGAAAGTAATCCCTCCTCACCTCCAGGAACTAGTCTYA	4020
Qy	4021	GCTRGCAAGAACTAATAGCCCTCACTTGGGGCACTAGAAATTAAGGAGAAAAAGGGYAAA	4080
Db	4021	GCTRGCAAGAACTAATAGCCCTCACTTGGGGCACTAGAAATTAAGGAGAAAAAGGGYAAA	4080
Qy	4081	TATATATACAGACTCTTRATATGCTTACCTAGTCTGTCCTCATGMRGCAATATGSGAR	4140
Db	4081	TATATATACAGACTCTTRATATGCTTACCTAGTCTGTCCTCATGMRGCAATATGSGAR	4140
Qy	4141	AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCMACATCAGGAAGCCATTAGGAR	4200
Db	4141	AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCMACATCAGGAAGCCATTAGGAR	4200
Qy	4201	ATTATTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCTGGGTCAATCA	4260
Db	4201	ATTATTATYTGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCTGGGTCAATCA	4260
Qy	4261	NAAGGAAAGRAAGGGAATTAAGRGAAATTCAGGAAATTCAGGAAACCAATATTAAGGAAAGG	4320
Db	4261	NAAGGAAAGRAAGGGAATTAAGRGAAATTCAGGAAATTCAGGAAACCAATATTAAGGAAAGG	4320
Qy	4321	TGCAAGGAGGAGCCCTCCATTTAGAAATGCTTATTAACCTTCCCTTATAGTATAGGTAATCC	4380
Db	4321	TGCAAGGAGGAGCCCTCCATTTAGAAATGCTTATTAACCTTCCCTTATAGTATAGGTAATCC	4380
Qy	4381	CTTCCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACGAGG	4440
Db	4381	CTTCCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACGAGG	4440
Qy	4441	CAGTTTTCTCCCTCGGAGCGGTTAGCCACTGAGAGGAGAAATTAATCTTTTGGCTGCAAC	4500
Db	4441	CAGTTTTCTCCCTCGGAGCGGTTAGCCACTGAGAGGAGAAATTAATCTTTTGGCTGCAAC	4500
Qy	4501	TATCAATGGAATTTACTTTAAACCTTCAATCAACCTTTCCTTACCTAGGATCGATAGCAC	4560
Db	4501	TATCAATGGAATTTACTTTAAACCTTCAATCAACCTTTCCTTACCTAGGATCGATAGCAC	4560
Qy	4561	CCATCARATGGCAATCAATTTTACTGGAACAGGCGCTTTTCAAACTTATCAAGCARAT	4620
Db	4561	CCATCARATGGCAATCAATTTTACTGGAACAGGCGCTTTTCAAACTTATCAAGCARAT	4620
Qy	4621	AKTCHAGGCTGTGAATGTGCGARAAATTAATCCCTGCTCTYATCGCCAAAGCTCTTTC	4680
Db	4621	AKTCHAGGCTGTGAATGTGCGARAAATTAATCCCTGCTCTYATCGCCAAAGCTCTTTC	4680
Qy	4681	AGGARAACAAAPAAACAGGCAATTCCTGGRAPAACTGGCACTGTATTTTACCCCAAG	4740
Db	4681	AGGARAACAAAPAAACAGGCAATTCCTGGRAPAACTGGCACTGTATTTTACCCCAAG	4740
Qy	4741	CCCAAACTCAGGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTACCGGTTGGGCA	4800
Db	4741	CCCAAACTCAGGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTACCGGTTGGGCA	4800

QY	4801	RAGSCCTTCCCTGTAGGACAGAAAAGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAA	4860
DB	4801		
QY	4861	ATAATTCAGATTCGAGCTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTCCAG	4920
DB	4861		
QY	4921	GCACAGTAACCCAGGAGATCCAGCGGTTAGGTATACGATACATCTTACACTGCGCC	4980
DB	4921		
QY	4981	TGAAGGCCACAGTCCTCAGGGAAGGTGAGAAAATGAATGAAYACTCAAAGGACATCTA	5040
DB	4981		
QY	5041	AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGTCGTGGCTATAGCCTTAAA	5100
DB	5041		
QY	5101	AGAACTGCAACTTTCACCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG	5160
DB	5101		
QY	5161	CCCTTCATAACCAATGACCTTGTGCTTGACCCAAAGACAGCCAACTTAGTTGAGACATCA	5220
DB	5161		
QY	5221	CCTCCTTAGCCAAATATCAAGCTTCTTAAACATTTACAGGAACCTATCCCTGAGAG	5280
DB	5221		
QY	5281	AGGAAAAGAACTATTCACCCWGTGACATGTTAGTCAAGTCCCTTCYCTCTAATT	5340
DB	5281		
QY	5341	CCCCATCCCTAGATACATCTCTGGGAAGGACCTACCCAGTCAATTTATYTAACCCCACTG	5400
DB	5341		
QY	5401	CGGTAAAGTGGCTGGAGTGAGCTTGGATACATCACTTGATCAATCTCGGATAC	5460
DB	5401		
QY	5461	TGCCAAAGAACCTGAAAATCCAGGACACGCTAGCTATCTCTGGAACCTCTAGAGG	5520
DB	5461		
QY	5521	ATTTGCGCTGCTCTTCAAAACAAACAGGAGGAAGTAACATAAAATCATAAATCCCCC	5580
DB	5521		
QY	5581	ATGSGCTCCCTTATCATATTTTCTCTKTASTGTTTATYACCTTTTCACTCTCACT	5640
DB	5581		
QY	5641	GCACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA	5700
DB	5641		
QY	5701	ATGCAGGCTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCC	5760
DB	5701		
QY	5761	ACCTTCACTGCCCAACCCATATGCCCCGCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
DB	5761		
QY	5821	CATGCAATACCTCATTTATGGACAGGAAAATGATTAATCTTAGTTGCTCGGAGGACTT	5880
DB	5821		
QY	5881	GGAGTCACTGTCTGTGGACTTATCTTCAACCCAACTGGTATGTCTGATGGGGGTGGAGTT	5940
DB	5881		

DB	5881	GGAGTCACTGTCTGTGGACTTATCTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTT	5940
QY	5941	CAAGATCAGGCAAGAGAAAACATGTAAAAGAGTAATCTCCCAACTCACCSSGGGTACAT	6000
DB	5941		
QY	6001	GGCACCTCTAGCCCTCACAAGGACTAGATCTCTCAAAACTACATGAAAACCTCCGCTACC	6060
DB	6001		
QY	6061	CATACCTGCTGTAGACCTTATTAATACCACTCTGCTGGCTCCATGAGGTCTCGGCC	6120
DB	6061		
QY	6121	CAAAACCTACTAACTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATC	6180
DB	6121		
QY	6181	CCTGTACTGAACAATGGAACAJCTTCAGACACAGAAATAAACCACTTCCGTTTTAGTA	6240
DB	6181		
QY	6241	GGACCTCTGTTTCCAAATSTGGAAATACCCATACCTCAAACTCAGCTGTGTAATTT	6300
DB	6241		
QY	6301	AGCAATACTACATACACAACCAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA	6360
DB	6301		
QY	6361	ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATTCGTTGTTG	6420
DB	6361		
QY	6421	AATGGCTCTCAGAACTATGTGCTTCTCATCTTCTAGTGCCCTCCATGRCCTATC	6480
DB	6421		
QY	6481	ACTGMAACAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCATT	6540
DB	6481		
QY	6541	CTTCTTTTGTATPAGGACAGAGTGCTAGGTGCACTAGTACTGTGCAATTTGGCGGTATC	6600
DB	6541		
QY	6601	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTTAAATGGGACATGGAACGG	6660
DB	6601		
QY	6661	GTGCGGACTCCCTGGTCACTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	6720
DB	6661		
QY	6721	CRAAATCGAAGAGCTTTAGACTYGTCTAACCGCTGAAGAGGGGAAACCTGTTTATTTTA	6780
DB	6721		
QY	6781	GGGGAAGAATGCTGTTTATTTATTTAACTCCGGAATCGTCACTGAGAAAGTTTAAAGAA	6840
DB	6781		
QY	6841	ATTCSAGATCGAATACAACTAGACAGAGAGCTTCGAAACACTGGACCTCGGGGCTC	6900
DB	6841		
QY	6901	CTCAGCCRAATGATGCTGATTCCTCCCTTTTGGAGCTCTAGCAGCTATATATTTG	6960
DB	6901		
QY	6961	CTACTCTCTTTGGACCTCTGATCTTTTACCTCTGTTTAACTTTGCTCTTCCAGAAATC	7020
DB	6961		

6961 CTACTCTCTTTGGACCCCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTCCAGAAATC 7020  
7021 GAAGCTGTRAACTTACAAATGGAGCCCAAGATGCAGTCCAGACTAAGATCTACCGCAGA 7080  
7021 GAAGCTGTRAACTTACAAATGGAGCCCAAGATGCAGTCCAGACTAAGATCTACCGCAGA 7080  
7081 CCCCTGACCGGGCTGTAGCCACGATCTGTATGATGTTAATGACATCAAAAGGCACCCCTCCT 7140  
7081 CCCCTGACCGGGCTGTAGCCACGATCTGTATGATGTTAATGACATCAAAAGGCACCCCTCCT 7140  
7141 GAGGAATCTCAGCTGCACAACTCTACTAGCCGCCCAATTCAGCAGGAAGCAGTTAGAGC 7200  
7141 GAGGAATCTCAGCTGCACAACTCTACTAGCCGCCCAATTCAGCAGGAAGCAGTTAGAGC 7200  
7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGACTGAGA 7260  
7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGACTGAGA 7260  
7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTTAGSTGGGAAGGTG 7320  
7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTTAGSTGGGAAGGTG 7320  
7321 ACCACATCCACCTTTAAACACGCGGCTTGCACTTAGTGTGACACCAATCAGAGAG 7380  
7321 ACCACATCCACCTTTAAACACGCGGCTTGCACTTAGTGTGACACCAATCAGAGAG 7380  
7381 CTCACATAAATGCTTAATTAGGCAAGACAGGAGTAAGAATAAGCAATCATYTTATTC 7440  
7381 CTCACATAAATGCTTAATTAGGCAAGACAGGAGTAAGAATAAGCAATCATYTTATTC 7440  
7441 MTGAGAGCAGCAGCAGGAGGACATGATCGGGATATAAAACCAAGTTCGAGCCGGCAA 7500  
7441 MTGAGAGCAGCAGCAGGAGGACATGATCGGGATATAAAACCAAGTTCGAGCCGGCAA 7500  
7501 CGGCAACCCCTTTGGGTCCCTCCCTTGTATGGAGCTCTGTTTCATGCTATTTTAC 7560  
7501 CGGCAACCCCTTTGGGTCCCTCCCTTGTATGGAGCTCTGTTTCATGCTATTTTAC 7560  
7561 TCTATTAAATCTTGCACTGCR 7582  
7561 TCTATTAAATCTTGCACTGCR 7582

RESULT 3  
AX000966  
LOCUS AX000966 7582 bp DNA linear PAT 10-MAR-2000  
DEFINITION Sequence 11 from Patent W09902696.  
ACCESSION AX000966  
VERSION AX000966.1 GI:7241208  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.  
REFERENCE 1 (bases 1 to 7582)  
AUTHORS Beseme,F. and Blond,J.  
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE  
DISEASES OR WITH PREGNANCY DISORDERS  
JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;  
BIO MERIEUX (FR); BESEME FREDERIC (FR)  
FEATURES  
source  
1. .7582  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 98.8%; Score 7489.6; DB 6; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
DB 1 CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTCTATTTAAATCTTGAACCTGCA 120  
DB 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTTCACTCTATTTAAATCTTGAACCTGCA 120  
QY 121 CTTCTTGCTGCTCAATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
DB 121 CTTCTTGCTGCTCAATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180  
QY 181 TGTGTTGCAACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 181 TGTGTTGCAACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 241 CGCTGTGCTCTGATCAGCAGAGCGCCCATTTGCGCTCTCCCAATTTGGGCTTAAAGGCTTGC 300  
DB 241 CGCTGTGCTCTGATCAGCAGAGCGCCCATTTGCGCTCTCCCAATTTGGGCTTAAAGGCTTGC 300  
QY 301 CATTTGNCCTGCAACCGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAAACACTANTCACT 360  
DB 301 CATTTGNCCTGCAACCGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAAACACTANTCACT 360  
QY 361 GGGTTCCATGGTTCCTTCTGTGACCCACGGCTTTCTAATAKAACTATAACACTTACCA 420  
DB 361 GGGTTCCATGGTTCCTTCTGTGACCCACGGCTTTCTAATAKAACTATAACACTTACCA 420  
QY 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480  
DB 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAATAC 480  
QY 481 GARGCTTGCCACCATCTTGGAAAGCGGCTGTCTACRRTCTTTGGAAGTGGTTTCCACCATC 540  
DB 481 GARGCTTGCCACCATCTTGGAAAGCGGCTGTCTACRRTCTTTGGAAGTGGTTTCCACCATC 540  
QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTTACATTTTGGCBACCCAGGACCGGACATCC 600  
DB 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTTACATTTTGGCBACCCAGGACCGGACATCC 600  
QY 601 MAAGTGATGGGAAACGTTCCCGCAAGACAAACCGCCCTAAGAGCTATTCTTGGARAAT 660  
DB 601 MAAGTGATGGGAAACGTTCCCGCAAGACAAACCGCCCTAAGAGCTATTCTTGGARAAT 660  
QY 661 TGGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCAGTGCC 720  
DB 661 TGGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCAGTGCC 720  
QY 721 GCCTGCACTCTGTGAGGAAAGTATAAATTTATACCATCTTACAGCTAGACATCTTTTG 780  
DB 721 GCCTGCACTCTGTGAGGAAAGTATAAATTTATACCATCTTACAGCTAGACATCTTTTG 780  
QY 781 TAGAAAGCAATTTGAGGTAAGTGCCATAAGTACAAACTTTCTTTTCTTATTAAGAGACAA 840  
DB 781 TAGAAAGCAATTTGAGGTAAGTGCCATAAGTACAAACTTTCTTTTCTTATTAAGAGACAA 840  
QY 841 CTCACAAATTTATGTAAGGTAAGTGTATTTATGCCCCCTACAGGAAGCTTTTACAGTCTAC 900  
DB 841 CTCACAAATTTATGTAAGGTAAGTGTATTTATGCCCCCTACAGGAAGCTTTTACAGTCTAC 900  
QY 901 CTATCCAGCACTCCCGACTCTTCTCCCAATTAAGGACCCCTTCAACCAATGG 960  
DB 901 CTATCCAGCACTCCCGACTCTTCTCCCAATTAAGGACCCCTTCAACCAATGG 960  
QY 961 TCCAAAGGAGATAGACAAAGGTAAGCAAGTGAACCAAGAGTGCCAAATTTTCCCAAT 1020  
DB 961 TCCAAAGGAGATAGACAAAGGTAAGCAAGTGAACCAAGAGTGCCAAATTTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGGAGTGCAATGTGCT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGGAGTGCAATGTGCT 1080  
QY 1081 TTTTCTCCAGACTTAAAGCAATTAAGGACAGCTTAGGTAAGTCTCAGATAATYCT 1140  
DB 1081 TTTTCTCCAGACTTAAAGCAATTAAGGACAGCTTAGGTAAGTCTCAGATAATYCT 1140

QY	1141	GATGGCTATATTGRTGTTTTCACAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAT	1200	Db	2221	GGCAGTCACTAGATATCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT	2280
Db	1141	GATGGCTATATTGRTGTTTTCACAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAT	1200	QY	2281	TCACATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTGTAGGGAGAGACATCTCA	2340
QY	1201	ATATATGCTACTGCTAAATCAGACATAACCCCAATAGAGAGAGTGCCACCACTACTGC	1260	Db	2281	TCACATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTGTAGGGAGAGACATCTCA	2340
Db	1201	ATATATGCTACTGCTAAATCAGACATAACCCCAATAGAGAGAGTGCCACCACTACTGC	1260	QY	2341	GCAAAAGCAGGGGCCATTATACACCTGAAACATAGGAGAGAGAACACCCGCTTTCTGTGNC	2400
QY	1261	AGCCTGAGRTGTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATPANGATGACAACA	1320	Db	2341	GCAAAAGCAGGGGCCATTATACACCTGAAACATAGGAGAGAGAACACCCGCTTTCTGTGNC	2400
Db	1261	AGCCTGAGRTGTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATPANGATGACAACA	1320	QY	2401	CTGTCTTGAGGAAGAAATTAATCTCTGAGTCTGGGCAACAGAAAGCAATATATGAGCAGC	2460
QY	1321	GAAGGAAGANAATGATTTCCCCACAGCCAGCARGCAGTTCCCGAGTCTASACCTCATTTG	1380	Db	2401	CTGTCTTGAGGAAGAAATTAATCTCTGAGTCTGGGCAACAGAAAGCAATATATGAGCAGC	2460
Db	1321	GAAGGAAGANAATGATTTCCCCACAGCCAGCARGCAGTTCCCGAGTCTASACCTCATTTG	1380	QY	2461	CAAAGAAATGCCGCTCTCTGTTCAAGTTAAACTAAAGAAATTCACCTCTCTTCCCTACCAA	2520
QY	1381	GGGACACAGAAATCAGTAAACATGGGAGATTGGTCTGTCAGACATTTGCTAACTTGTGTG	1440	Db	2461	CAAAGAAATGCCGCTCTCTGTTCAAGTTAAACTAAAGAAATTCACCTCTCTTCCCTACCAA	2520
Db	1381	GGGACACAGAAATCAGTAAACATGGGAGATTGGTCTGTCAGACATTTGCTAACTTGTGTG	1440	QY	2521	GGCAGTACCCCTCTCAGACCCCAAGGCCCAACAGGAATTCCAAAGAAATTTGTTAAGGACTTAA	2580
QY	1441	TASAGGACTAAGGAACCTTASGAAGAAATCTAYGAATTAATCAATGATGTCACACATA	1500	Db	2521	GGCAGTACCCCTCTCAGACCCCAAGGCCCAACAGGAATTCCAAAGAAATTTGTTAAGGACTTAA	2580
Db	1441	TASAGGACTAAGGAACCTTASGAAGAAATCTAYGAATTAATCAATGATGTCACACATA	1500	QY	2581	AAGCCCAAGGCTTAGTAAACCAATGCAATCTCCCTGCAAGTAAATCCGTAGTGGATTGAG	2640
QY	1501	ACACAGGGGAAGGAAGAAAATCTACTGCTTTCTGAGAGACTAAGGGAGCAATTGAG	1560	Db	2581	AAGCCCAAGGCTTAGTAAACCAATGCAATCTCCCTGCAAGTAAATCCGTAGTGGATTGAG	2640
Db	1501	ACACAGGGGAAGGAAGAAAATCTACTGCTTTCTGAGAGACTAAGGGAGCAATTGAG	1560	QY	2641	GAGGACAGAAACCCAGTGGACAGAGTGGAGGTTAGTGCAGATCTCAGGATTAATCAATGG	2700
QY	1561	GAAGCGTCCCTCTCTGTCACCTGACTCTTCTGAAGCCCACTAATCTTAAAGCGTAAGTT	1620	Db	2641	GAGGACAGAAACCCAGTGGACAGAGTGGAGGTTAGTGCAGATCTCAGGATTAATCAATGG	2700
Db	1561	GAAGCGTCCCTCTCTGTCACCTGACTCTTCTGAAGCCCACTAATCTTAAAGCGTAAGTT	1620	QY	2701	AGGCGGTGTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATAC	2760
QY	1621	TATCACTCAGTCAAGTGCAGACATTAGAAAACTTCAAAAGTCTGCGTAGGCCCGGAG	1680	Db	2701	AGGCGGTGTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATAC	2760
Db	1621	TATCACTCAGTCAAGTGCAGACATTAGAAAACTTCAAAAGTCTGCGTAGGCCCGGAG	1680	QY	2761	CAGAGGAGCAGAGTGGTTTACASTCTGGAACCTTGAAGATGCTTCTGATCCCTG	2820
QY	1681	CAAAAATTAGAAACCCCTATTGAACCTTGGCAACTTCGGTTTTTTTATATAGAGATCAGGAG	1740	Db	2761	CAGAGGAGCAGAGTGGTTTACASTCTGGAACCTTGAAGATGCTTCTGATCCCTG	2820
Db	1681	CAAAAATTAGAAACCCCTATTGAACCTTGGCAACTTCGGTTTTTTTATATAGAGATCAGGAG	1740	QY	2821	TACATCTGACTCTCAATCTCTTGTGCTTTCAGAGATCTTCAAAACCCARCATCTCAAC	2880
QY	1741	GAGCAGCGGAAACAGGACAAACCGGATTAATAAGGCAACCGCTTTAGTCATGACCT	1800	Db	2821	TACATCTGACTCTCAATCTCTTGTGCTTTCAGAGATCTTCAAAACCCARCATCTCAAC	2880
Db	1741	GAGCAGCGGAAACAGGACAAACCGGATTAATAAGGCAACCGCTTTAGTCATGACCT	1800	QY	2881	TCACCTGGACTTTTACCCCAAGGGTTCAAGGATAGYCCCATCTATTTGGCCAGGCAT	2940
QY	1801	CAGCAAGTGCATTTGAGGCTCTGGAAGGGGAAAGCTGGGCAATTTGAATGGCTTAA	1860	Db	2881	TCACCTGGACTTTTACCCCAAGGGTTCAAGGATAGYCCCATCTATTTGGCCAGGCAT	2940
Db	1801	CAGCAAGTGCATTTGAGGCTCTGGAAGGGGAAAGCTGGGCAATTTGAATGGCTTAA	1860	QY	2941	TAGCCCAAGACTTTGAGYCATYMTCTACTCTGACACTCTTGTCTTCRGTAAGTGGATG	3000
QY	1861	TAGGGCTTGCTTCCAGTCCGCTCTACAAGGACACTTTTAAAAAGATTGTCCAAGTAGAAG	1920	Db	2941	TAGCCCAAGACTTTGAGYCATYMTCTACTCTGACACTCTTGTCTTCRGTAAGTGGATG	3000
Db	1861	TAGGGCTTGCTTCCAGTCCGCTCTACAAGGACACTTTTAAAAAGATTGTCCAAGTAGAAG	1920	QY	3001	ATTTTACTTTTTRGCGCYRTTCAAGAACCTTGTGCCATCAAGCCACCAAGCCTCTTMA	3060
QY	1921	TAGAGCCGCCCTTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGGAAGGCCCTGCC	1980	Db	3001	ATTTTACTTTTTRGCGCYRTTCAAGAACCTTGTGCCATCAAGCCACCAAGCCTCTTMA	3060
Db	1921	TAGAGCCGCCCTTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGGAAGGCCCTGCC	1980	QY	3061	ATTTCTCTCGYACTGTGGCTACAWGGTTTCCAAACSAARARGCTCARCTCTGTCTCAGC	3120
QY	1981	CAGGGHCAAGGCTTTTTCAGTCAGAGCCACTTAAACCATGATTCAGCAGCAGGACTG	2040	Db	3061	ATTTCTCTCGYACTGTGGCTACAWGGTTTCCAAACSAARARGCTCARCTCTGTCTCAGC	3120
Db	1981	CAGGGHCAAGGCTTTTTCAGTCAGAGCCACTTAAACCATGATTCAGCAGCAGGACTG	2040	QY	3121	AGGTTAAATCTTAGGCTAARATTTATCCAAAGGCACCARGGCCCTCAGTGAGGAARYA	3180
QY	2041	AGGTTGCTTGGGCAAGCGCCATCCCATGCCATCAACCTCTCACAGAGCCCTGGGTATGCTT	2100	Db	3121	AGGTTAAATCTTAGGCTAARATTTATCCAAAGGCACCARGGCCCTCAGTGAGGAARYA	3180
Db	2041	AGGTTGCTTGGGCAAGCGCCATCCCATGCCATCAACCTCTCACAGAGCCCTGGGTATGCTT	2100	QY	3181	TCCAGCCTTACTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTTG	3240
QY	2101	GACATTTGAGGCGCAGGAAGGTTGTCCTCTGGAACACTGGTGGCTCTTTAGTCTTACT	2160	Db	3181	TCCAGCCTTACTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTTG	3240
Db	2101	GACATTTGAGGCGCAGGAAGGTTGTCCTCTGGAACACTGGTGGCTCTTTAGTCTTACT	2160	QY	3241	GCTTAAVAGGYTTCTGCGGAATGGAATTCCTCCAGGTTGTCRAAATAGCCAGYCATTA	3300
QY	2161	CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGCG	2220	Db	3241	GCTTAAVAGGYTTCTGCGGAATGGAATTCCTCCAGGTTGTCRAAATAGCCAGYCATTA	3300
Db	2161	CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGCG	2220	QY	3301	WATACASTAATTAAGGAAACTCAGAAAGCCATATCCCATTTTARTTAAGATGGAYAMCTGAA	3360
QY	2221	GGCAGTCACTAGATACTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT	2280				



Db	3301	WATACASTAATTAAGAAACTCAGAAAGCCATATCCCATTTARTAGATGGAYAMCTGAA	3360
Qy	3361	GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGCCCTTAAACCCAGYVCCAGTGTAAAGYT	3420
Db	3361	GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGCCCTTAAACCCAGYVCCAGTGTAAAGYT	3420
Qy	3421	TGCCAACHGGGCAAGACTTTTSTYATYRTCAAGAAAAAACAAGAAAGCTCTRGGA	3480
Db	3421	TGCCAACHGGGCAAGACTTTTSTYATYRTCAAGAAAAAACAAGAAAGCTCTRGGA	3480
Qy	3481	GTCTCTTACACAGRTCCRAGGAGYAGCTTGCACCCYRTGGCRYACCTGASTAAGAAAYT	3540
Db	3481	GTCTCTTACACAGRTCCRAGGAGYAGCTTGCACCCYRTGGCRYACCTGASTAAGAAAYT	3540
Qy	3541	GATGTAGTGGCAAGGGTGTGRCTTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA	3600
Db	3541	GATGTAGTGGCAAGGGTGTGRCTTCATTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA	3600
Qy	3601	GTATCTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Db	3601	GTATCTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Qy	3661	GTGAAYRGCACTACTCTGTCTTAAGGAGACTTGTGGCTGTGCAGAACACYGTGTACTTAAA	3720
Db	3661	GTGAAYRGCACTACTCTGTCTTAAGGAGACTTGTGGCTGTGCAGAACACYGTGTACTTAAA	3720
Qy	3721	TRTCAGGCTCTATTAATGAAGGCCAGTGTCTGCRACCTGTGCACTGTGCAACTCTTAAC	3780
Db	3721	TRTCAGGCTCTATTAATGAAGGCCAGTGTCTGCRACCTGTGCACTGTGCAACTCTTAAC	3780
Qy	3781	CCAGYCNCAATTTCTTCAGACAATGAAGAAAAAGATARAAYATACTGTCAACAARTTAAT	3840
Db	3781	CCAGYCNCAATTTCTTCAGACAATGAAGAAAAAGATARAAYATACTGTCAACAARTTAAT	3840
Qy	3841	TCTCAACCTACTGCACTCAGGGGACCTTCTAGAGTTCCYTTGACTGATCCYGCCTT	3900
Db	3841	TCTCAACCTACTGCACTCAGGGGACCTTCTAGAGTTCCYTTGACTGATCCYGCCTT	3900
Qy	3901	CAACTTGTATCTAGTGAAGTTCTTTGTAGAAAAAGACTTTCGAAAAAGGGGTATGC	3960
Db	3901	CAACTTGTATCTAGTGAAGTTCTTTGTAGAAAAAGACTTTCGAAAAAGGGGTATGC	3960
Qy	3961	AGTGGTCACTGATPAATGGAATAYTTGAAAGTAATCCCTCCTCAGGAACTAGTGCTVA	4020
Db	3961	AGTGGTCACTGATPAATGGAATAYTTGAAAGTAATCCCTCCTCAGGAACTAGTGCTVA	4020
Qy	4021	GCTRGCAAGTAATAGCCCTTCAITVGGGCACTAGAAATTAGGAGAAAGAAAAAGGYAAA	4080
Db	4021	GCTRGCAAGTAATAGCCCTTCAITVGGGCACTAGAAATTAGGAGAAAGAAAAAGGYAAA	4080
Qy	4081	TATATATACAGACTCTRAATATGCTVACCTAGTTCNTCCATGCCATGMRGCAATAGSAR	4140
Db	4081	TATATATACAGACTCTRAATATGCTVACCTAGTTCNTCCATGCCATGMRGCAATAGSAR	4140
Qy	4141	AGAAAGGGAATTCCTAATCTCYGAGGAACACCTATCMACATCAGGAAGCCATTAGGAR	4200
Db	4141	AGAAAGGGAATTCCTAATCTCYGAGGAACACCTATCMACATCAGGAAGCCATTAGGAR	4200
Qy	4201	ATTATATYTGCGWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTCYGGGGTCATCA	4260
Db	4201	ATTATATYTGCGWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTCYGGGGTCATCA	4260
Qy	4261	NAAAGGAAGRAAGGGAATASAGRAAYTGCCNAGCAKATATTGAAGCMAAAAGAGC	4320
Db	4261	NAAAGGAAGRAAGGGAATASAGRAAYTGCCNAGCAKATATTGAAGCMAAAAGAGC	4320
Qy	4321	TGCAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTCCCTAGTATAGGGTAATCC	4380
Db	4321	TGCAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTCCCTAGTATAGGGTAATCC	4380
Qy	4381	CTTCGGGGAACCAAGCCCACTACTACAGCAGGAGAAACAGAAATGGGGAACCTCAGGAGG	4440
Db	4381	CTTCGGGGAACCAAGCCCACTACTACAGCAGGAGAAACAGAAATGGGGAACCTCAGGAGG	4440

Qy	4441	CAGTTTCTCCCTCGGGAAGCGTTAGCCACTGAAGAGGGGAAAAATCTTTTGCCTGCAAC	4500
Db	4441	CAGTTTCTCCCTCGGGAAGCGTTAGCCACTGAAGAGGGGAAAAATCTTTTGCCTGCAAC	4500
Qy	4501	TATCAATGGAATTTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATAGCAC	4560
Db	4501	TATCAATGGAATTTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATAGCAC	4560
Qy	4561	CCATCARATGGCCAAATCATTTATTTACTGGACAGGCTTTTCAAAACTATCAAGCARAT	4620
Db	4561	CCATCARATGGCCAAATCATTTATTTACTGGACAGGCTTTTCAAAACTATCAAGCARAT	4620
Qy	4621	AKTCAGGGCTGTGAATGTGCCARAPAAATAATATCCCTGCTCTATCGCCAAAGCTCTTC	4680
Db	4621	AKTCAGGGCTGTGAATGTGCCARAPAAATAATATCCCTGCTCTATCGCCAAAGCTCTTC	4680
Qy	4681	AGGARAACAARAACAGGCCATTTACCTGRARARARACTGGCAACTGATTTTACCACAAG	4740
Db	4681	AGGARAACAARAACAGGCCATTTACCTGRARARARACTGGCAACTGATTTTACCACAAG	4740
Qy	4741	CCCAACCTCAGGGATTTTCACTAGTCTAGTCTGGTARATATCTTTTCAAGGCTTGGCA	4800
Db	4741	CCCAACCTCAGGGATTTTCACTAGTCTAGTCTGGTARATATCTTTTCAAGGCTTGGCA	4800
Qy	4801	RAGGCCCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCAAG	4860
Db	4801	RAGGCCCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCAAG	4860
Qy	4861	ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATGCTTGTTCAG	4920
Db	4861	ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATGCTTGTTCAG	4920
Qy	4921	GCCACAGTAAACCCAGGGAGTATCCAGAGGCTTACAGAGTGACAAATGCTTGTTCAG	4980
Db	4921	GCCACAGTAAACCCAGGGAGTATCCAGAGGCTTACAGAGTGACAAATGCTTGTTCAG	4980
Qy	4981	TGAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAATGAATGAAYACTCAAGGACATCTA	5040
Db	4981	TGAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAATGAATGAAYACTCAAGGACATCTA	5040
Qy	5041	AAAAAGCAACCCAGGAAACCCACCTCAGATGGCTGTCTGTGCTATAGGCTTAAAA	5100
Db	5041	AAAAAGCAACCCAGGAAACCCACCTCAGATGGCTGTCTGTGCTATAGGCTTAAAA	5100
Qy	5101	AGAACTCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAG	5160
Db	5101	AGAACTCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAG	5160
Qy	5161	CCCTTCAATAACCAATGACTTGTGCTTGAACCAAGACGCAACTTAGTTCAGACATCA	5220
Db	5161	CCCTTCAATAACCAATGACTTGTGCTTGAACCAAGACGCAACTTAGTTCAGACATCA	5220
Qy	5221	CCTCTTAGCCAAATATCAACAGTCTTAAACATTAACAGGAACCTTCCCTGAGAAG	5280
Db	5221	CCTCTTAGCCAAATATCAACAGTCTTAAACATTAACAGGAACCTTCCCTGAGAAG	5280
Qy	5281	AGGAAAAAGAACTTATCCACCCWGTGACATGTTAGTCAAGTCCCTTCTCTAAT	5340
Db	5281	AGGAAAAAGAACTTATCCACCCWGTGACATGTTAGTCAAGTCCCTTCTCTAAT	5340
Qy	5341	CCCAATCTCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATVATCCCACTG	5400
Db	5341	CCCAATCTCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATVATCCCACTG	5400
Qy	5401	CGGTTAAGTGGCTGGAGTGGATCTTGGATACATCACTTGAAGTCAAACTCCTGGATAC	5460
Db	5401	CGGTTAAGTGGCTGGAGTGGATCTTGGATACATCACTTGAAGTCAAACTCCTGGATAC	5460
Qy	5461	TGCCAAAGGAACCTTGAAAAATCCAGGAGCAACGCTAGCTATTTCTGTGCAACCTCTAGAGG	5520
Db	5461	TGCCAAAGGAACCTTGAAAAATCCAGGAGCAACGCTAGCTATTTCTGTGCAACCTCTAGAGG	5520

QY	5521	ATTGCGCCTGCTCTTCAAAACAAACAGAGGAAAGTAACATAAAATCAATAATCCCCC	5580
DB	5521	ATTGCGCCTGCTCTTCAAAACAAACAGAGGAAAGTAACATAAAATCAATAATCCCCC	5580
QY	5581	ATGSGCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTCTTTCACCTCACT	5640
DB	5581	ATGSGCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTCTTTCACCTCACT	5640
QY	5641	GCACCCCTCCATCCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGAGA	5700
DB	5641	GCACCCCTCCATCCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGAGA	5700
QY	5701	ATGACGGCTCCGGAANAATGATGCCCCATGCTATAGGAGTCTTTTAAAGGGAACCCCC	5760
DB	5701	ATGACGGCTCCGGAANAATGATGCCCCATGCTATAGGAGTCTTTTAAAGGGAACCCCC	5760
QY	5761	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
DB	5761	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
QY	5821	CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCTAGTGTCTCTGGAGGACTT	5880
DB	5821	CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCTAGTGTCTCTGGAGGACTT	5880
QY	5881	GGAGTCACTGCTGTGGACTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTT	5940
DB	5881	GGAGTCACTGCTGTGGACTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTT	5940
QY	5941	CAAGATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCACTCACCGGGTACAT	6000
DB	5941	CAAGATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCACTCACCGGGTACAT	6000
QY	6001	GGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	6060
DB	6001	GGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	6060
QY	6061	CATACCTGCTGTAAGCCTATTTAATAACCACTCTCACTGGGCTCCATGAGGTTCTGGCC	6120
DB	6061	CATACCTGCTGTAAGCCTATTTAATAACCACTCTCACTGGGCTCCATGAGGTTCTGGCC	6120
QY	6121	CAAAACCTCTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTCAATC	6180
DB	6121	CAAAACCTCTAATCTGTGGATATGCTCCCTCCCTGAACTTCARGGCCATATGTTCAATC	6180
QY	6181	CCTGTACTGAAACAACTGTAAGCAACCTTACAGCAGAAATAAACAACACCTTCGGTTTAGTA	6240
DB	6181	CCTGTACTGAAACAACTGTAAGCAACCTTACAGCAGAAATAAACAACACCTTCGGTTTAGTA	6240
QY	6241	GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCACTCACTCACTGTTGTAATAATTT	6300
DB	6241	GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCACTCACTCACTGTTGTAATAATTT	6300
QY	6301	AGCAATCTACATACAAACCACTCCCAATGATCAGGTGGTAACTCTCCACACAA	6360
DB	6301	AGCAATCTACATACAAACCACTCCCAATGATCAGGTGGTAACTCTCCACACAA	6360
QY	6361	ATAGCTGCTTACCTCAGGAATAATTTTCTGTGTGTGCTACCTCAGCCCTATCGTTGTTG	6420
DB	6361	ATAGCTGCTTACCTCAGGAATAATTTTCTGTGTGTGCTACCTCAGCCCTATCGTTGTTG	6420
QY	6421	AATGGCTCTTCAGAACTATGCTTCTCTCATTTAGTGGCCCTCATGRCATCTTAC	6480
DB	6421	AATGGCTCTTCAGAACTATGCTTCTCTCATTTAGTGGCCCTCATGRCATCTTAC	6480
QY	6481	ACTGAACAAGATTATACAGTTATGCTATATCTAAGCCCGCAACAAAGAGTACCCATT	6540
DB	6481	ACTGAACAAGATTATACAGTTATGCTATATCTAAGCCCGCAACAAAGAGTACCCATT	6540
QY	6541	CTTCCCTTTGTTATAGGACGAGGAGTGTAGTGTACCTAGTGTGCTATGCGGCTATC	6600
DB	6541	CTTCCCTTTGTTATAGGACGAGGAGTGTAGTGTACCTAGTGTGCTATGCGGCTATC	6600
QY	6601	ACAACTCTACTCAGTTCTACTACAACTATCTCTCAAGAACTAAATGGGGACATGGAAACGG	6660

DB	6601	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGG	6660
QY	6661	GTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	6720
DB	6661	GTGCGGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	6720
QY	6721	CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGGAACCTGTTATTTTA	6780
DB	6721	CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGGAACCTGTTATTTTA	6780
QY	6781	GGGGAAGAATGCTGTTATTTATGTTATCAATCCGGATCGTCACTGAGAAAGTTTAAAGAA	6840
DB	6781	GGGGAAGAATGCTGTTATTTATGTTATCAATCCGGATCGTCACTGAGAAAGTTTAAAGAA	6840
QY	6841	ATTCSAGATCGAATACAAAGTACAGAGAGAGCTTCGAAACACTGGACCCCTGGGGCTC	6900
DB	6841	ATTCSAGATCGAATACAAAGTACAGAGAGAGCTTCGAAACACTGGACCCCTGGGGCTC	6900
QY	6901	CTCAGCCRAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
DB	6901	CTCAGCCRAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
QY	6961	CTACTCTCTTTGGAACCTGTATCTTTTACCTCTGTTTAACTTTGTCTCTCCAGAACTC	7020
DB	6961	CTACTCTCTTTGGAACCTGTATCTTTTACCTCTGTTTAACTTTGTCTCTCCAGAACTC	7020
QY	7021	GAAGCTGTAAATCTACAAATGGAGGCCCAAGATGCAGTCCAAAGACTAAGATCTACCGCAGA	7080
DB	7021	GAAGCTGTAAATCTACAAATGGAGGCCCAAGATGCAGTCCAAAGACTAAGATCTACCGCAGA	7080
QY	7081	CCCTTGGACCGGCTGTGTAGCCCAAGATCTGATGTTTAAATGACATCAAAGGCACCCCTCT	7140
DB	7081	CCCTTGGACCGGCTGTGTAGCCCAAGATCTGATGTTTAAATGACATCAAAGGCACCCCTCT	7140
QY	7141	GAGAAATCTCAGCTGACAAACCTCTACTAGCCGCCAAATTCAGCAGAGCAGTTAGAGC	7200
DB	7141	GAGAAATCTCAGCTGACAAACCTCTACTAGCCGCCAAATTCAGCAGAGCAGTTAGAGC	7200
QY	7201	GTTGTCGGCCAACTCCCAACAGCACTTGTAGTTTCTGTTGAGATGGGGACTGAGA	7260
DB	7201	GTTGTCGGCCAACTCCCAACAGCACTTGTAGTTTCTGTTGAGATGGGGACTGAGA	7260
QY	7261	GACAGCTAGCTGGAATTTCTTAGGCTGATTAAGAATCCVTAAGCCTAGTGGGAAAGTG	7320
DB	7261	GACAGCTAGCTGGAATTTCTTAGGCTGATTAAGAATCCVTAAGCCTAGTGGGAAAGTG	7320
QY	7321	ACCACTCCAGCTTTAAACAGGGGCTTGCNACTTAGTTCACACTGACCAATCAGAGAG	7380
DB	7321	ACCACTCCAGCTTTAAACAGGGGCTTGCNACTTAGTTCACACTGACCAATCAGAGAG	7380
QY	7381	CTCACTAAATGCTTAATTTAGGCAAGAGCAGAGGTAAAGAAATAGCCAATCATTTATTC	7440
DB	7381	CTCACTAAATGCTTAATTTAGGCAAGAGCAGAGGTAAAGAAATAGCCAATCATTTATTC	7440
QY	7441	MTGAGACACAGCAGGAGGAGCAATGATCGGGATATAAACCAGTTCGAGCCGCGCAA	7500
DB	7441	MTGAGACACAGCAGGAGGAGCAATGATCGGGATATAAACCAGTTCGAGCCGCGCAA	7500
QY	7501	CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC	7560
DB	7501	CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC	7560
QY	7561	TCTATTAAATCTGCACTGCR 7582	
DB	7561	TCTATTAAATCTGCACTGCR 7582	

RESULT 4					
AX027480					
LOCUS	AX027480	7582 bp	DNA	linear	PAT 16-SEP-2000
DEFINITION	Sequence 30 from Patent FR2788784.				
ACCESSION	AX027480				

VERSION	AX027480.1	GI:10188444
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 Mallet, F., Voisset, C. and Paranhos, B.G.	
AUTHORS	Patent: FR 2788784-A 30 28-JUL-2000;	
JOURNAL	BIO MERIEUX (FR)	
FEATURES	Location/Qualifiers	
source	1..7582	
ORIGIN	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
Query Match	98.8%;	Score 7489.6; DB 6; Length 7582;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 7582;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	CAACAATCGGGATATAAACCCAGCATTTCAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
DB	1	CAACAATCGGGATATAAACCCAGCATTTCAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
QY	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCATCTATTAAATCTTGCAACTGCA 120
DB	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCATCTATTAAATCTTGCAACTGCA 120
QY	121	CTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTGTCACCGTCCACCACTGC 180
DB	121	CTCTTCGTGTCATGTTTCTTACGGCTCGAGCTGAGCTTTGTCACCGTCCACCACTGC 180
QY	181	TGTTTCCACACCGCANACTCGCGTGACTCCCATCCCTCTGGATCTCTGAGGTGTC 240
DB	181	TGTTTCCACACCGCANACTCGCGTGACTCCCATCCCTCTGGATCTCTGAGGTGTC 240
QY	241	CGCTGTGCTCTGATCAGCGARGCGCCATTCGCGCTCCCAATTCGGCTTAAAGGCTTCG 300
DB	241	CGCTGTGCTCTGATCAGCGARGCGCCATTCGCGCTCCCAATTCGGCTTAAAGGCTTCG 300
QY	301	CATTGTNCTGCACGGCTAAGTGCCTGGGTTTCTTAATGTAGCTGAACACTANTCACT 360
DB	301	CATTGTNCTGCACGGCTAAGTGCCTGGGTTTCTTAATGTAGCTGAACACTANTCACT 360
QY	361	GGGTTCCATGTTCTTCTGTGACCCACCGGCTTCTAATAKAACTATAACACTTACCACA 420
DB	361	GGGTTCCATGTTCTTCTGTGACCCACCGGCTTCTAATAKAACTATAACACTTACCACA 420
QY	421	TGGCCCAAGATTCCATTCTTGGATCCGTGAGGSCAAGAACTCCAGGTCAGAGATAC 480
DB	421	TGGCCCAAGATTCCATTCTTGGATCCGTGAGGSCAAGAACTCCAGGTCAGAGATAC 480
QY	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCRCTTTTGGAAAGTGTTCACCAACCATC 540
DB	481	GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCRCTTTTGGAAAGTGTTCACCAACCATC 540
QY	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAAGGACATCC 600
DB	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAAGGACATCC 600
QY	601	MAAGTGTGGGAAACGTTTCCCGCAAGACAAACCGCCCTAAGACGTTATCTGGARAAT 660
DB	601	MAAGTGTGGGAAACGTTTCCCGCAAGACAAACCGCCCTAAGACGTTATCTGGARAAT 660
QY	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTGCAAGTGC 720
DB	661	TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCTGCAAGTGC 720
QY	721	GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTTTACAGCTAGACCTTTTGT 780
DB	721	GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTTTACAGCTAGACCTTTTGT 780
QY	781	TAGAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGCAA 840
DB	781	TAGAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGCAA 840
QY	841	CTCACAATATTGTAAGAAAGTGTGATTTATCCCTTACAGGAGGCTTCCAGAGTCTACCTCC 900
DB	841	CTCACAATATTGTAAGAAAGTGTGATTTATCCCTTACAGGAGGCTTCCAGAGTCTACCTCC 900
QY	901	CTATCCAGCATCCCGGACTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960
DB	901	CTATCCAGCATCCCGGACTCTTCCCAATTAAGGACCCCTTCAACCCCAATGG 960
QY	961	TCCAAAGGAGTAGACAAAGGGTAAACAGTGAAACAAAGAGTGCCCAATTTCCCCCAAT 1020
DB	961	TCCAAAGGAGTAGACAAAGGGTAAACAGTGAAACAAAGAGTGCCCAATTTCCCCCAAT 1020
QY	1021	TATGACCCCTTCCCAAGCAGTGGGAGGAAGAAATTCGGCCACCCAGAGTGCAATGTCYT 1080
DB	1021	TATGACCCCTTCCCAAGCAGTGGGAGGAAGAAATTCGGCCACCCAGAGTGCAATGTCYT 1080
QY	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAGACACTTAGGTAAATTTCTCAGATAATYCCCT 1140
DB	1081	TTTTTCTCCAGACTTAAAGCAAAATAAAGACACTTAGGTAAATTTCTCAGATAATYCCCT 1140
QY	1141	GATGGCTATATTGRTGTTTTTACAAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200
DB	1141	GATGGCTATATTGRTGTTTTTACAAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200
QY	1201	ATATATGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTCCACCACTAATCTGC 1260
DB	1201	ATATATGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTCCACCACTAATCTGC 1260
QY	1261	AGCCTGAGGTTTGGGATCTCTGATCTCAGTCAAGTCAATGGATANGGATGACAACA 1320
DB	1261	AGCCTGAGGTTTGGGATCTCTGATCTCAGTCAAGTCAATGGATANGGATGACAACA 1320
QY	1321	GAAGGAAAGANAATGATTTCCCAAGGCGAGCAGTCCAGTCTASACCTCATTTG 1380
DB	1321	GAAGGAAAGANAATGATTTCCCAAGGCGAGCAGTCCAGTCTASACCTCATTTG 1380
QY	1381	GGGACACAGAAATCAGTAACATGGGAGATTGGTGTGCGAGACATTTGCTTAACCTGTGTC 1440
DB	1381	GGGACACAGAAATCAGTAACATGGGAGATTGGTGTGCGAGACATTTGCTTAACCTGTGTC 1440
QY	1441	TASAAAGACTAAGGAAAATCTAGGAAAGAAATCTAYGAATTACTCAATGATGTCCACCAT 1500
DB	1441	TASAAAGACTAAGGAAAATCTAGGAAAGAAATCTAYGAATTACTCAATGATGTCCACCAT 1500
QY	1501	ACACAGGGGAAGGGAAGAAATCTCTGCTTTCTGGAGAGACTAAGGAGGCAATTTGAG 1560
DB	1501	ACACAGGGGAAGGGAAGAAATCTCTGCTTTCTGGAGAGACTAAGGAGGCAATTTGAG 1560
QY	1561	GAAGGCTGCTCTCTGTGTCACCTGACTCTTCTGAGGCGCAACTTAATCTTAAAGCGTAAGTT 1620
DB	1561	GAAGGCTGCTCTCTGTGTCACCTGACTCTTCTGAGGCGCAACTTAATCTTAAAGCGTAAGTT 1620
QY	1621	TATCACTCAGTCACTGAGCAGATTTAGAAAAAATCTTCAAAAGTCTGCCGTAGGCCGAG 1680
DB	1621	TATCACTCAGTCACTGAGCAGATTTAGAAAAAATCTTCAAAAGTCTGCCGTAGGCCGAG 1680
QY	1681	CABAACCTTAGAAAACCTATTGAACCTTGCAACCTTCGGTTTTTTTATAATAGATCAGGAG 1740
DB	1681	CABAACCTTAGAAAACCTATTGAACCTTGCAACCTTCGGTTTTTTTATAATAGATCAGGAG 1740
QY	1741	GAGCAGGCGGAACAGGACAAACCGGATTTAAAAAAGGCGCCACCGCTTTAGTCAACCTCT 1800
DB	1741	GAGCAGGCGGAACAGGACAAACCGGATTTAAAAAAGGCGCCACCGCTTTAGTCAACCTCT 1800
QY	1801	CAGGCAATGAGCTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATGGAATGCCTAA 1860
DB	1801	CAGGCAATGAGCTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATGGAATGCCTAA 1860
QY	1861	TAGGGCTTGCTTCCAGTGGGCTCTACAGGACACTTTTAAAAAGAGTTGCTCCAAAGTAGAG 1920

1861 TAGGGCTGCTCCAGTGGGCTCAACAAGACACTTTAAAAAGATTGTCACAGTAGAAG 1920  
1921 TAAGCGCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGCCCACTGCCC 1980  
1921 TAAGCGCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAGCCCACTGCCC 1980  
1981 CAGGGACAAGAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCAGCAGCAGACTG 2040  
1981 CAGGGACAAGAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCAGCAGCAGACTG 2040  
2041 AGGGTCCCTGGGCAAGCGCAATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTT 2100  
2041 AGGGTCCCTGGGCAAGCGCAATCCCATGCCATCAACCTCACAGAGCCCTGGGTATGCTT 2100  
2101 GACCATTTGAGGCGCAGGAAGTTGTCTCTGGGCACTGGTGGGCTCTTCTTAGTCTTACT 2160  
2101 GACCATTTGAGGCGCAGGAAGTTGTCTCTGGGCACTGGTGGGCTCTTCTTAGTCTTACT 2160  
2161 CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
2161 CTTCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGGAGACATCTTA 2340  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGGAGACATCTTA 2340  
2341 GCAAAAGAGCGGCGCAATATACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTNCC 2400  
2341 GCAAAAGAGCGGCGCAATATACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTNCC 2400  
2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGAGCAGC 2460  
2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGAGCAGC 2460  
2461 CAAAGAAATGCCGTCCTGTTCAAGTTAACTAAGGATTCACCTTCCTTCCCTACCAA 2520  
2461 CAAAGAAATGCCGTCCTGTTCAAGTTAACTAAGGATTCACCTTCCTTCCCTACCAA 2520  
2521 GGCAGTACCCCTCAGACCCAGCCCAACAGGATTCCAAAGATTTGTTAAGCACTTA 2580  
2521 GGCAGTACCCCTCAGACCCAGCCCAACAGGATTCCAAAGATTTGTTAAGCACTTA 2580  
2581 AAGCCCAAGGCTTAGTAAACCCATGCATACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
2581 AAGCCCAAGGCTTAGTAAACCCATGCATACTCCCTGCAGTAATTCGGTAGTGGATTGAG 2640  
2641 GAGGCAAGAAACCCAGTGAAGGAGGTTAGTCAAGATCTCAGGATTAATCAATGG 2700  
2641 GAGGCAAGAAACCCAGTGAAGGAGGTTAGTCAAGATCTCAGGATTAATCAATGG 2700  
2701 AGGCGGTGCTTTTATACCCAGTCTACTAGCCCTTACTGTGMYTTCCCAATAC 2760  
2701 AGGCGGTGCTTTTATACCCAGTCTACTAGCCCTTACTGTGMYTTCCCAATAC 2760  
2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAAGGATGCTTTCTGCAATCCCTG 2820  
2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAAGGATGCTTTCTGCAATCCCTG 2820  
2821 TACATCCTGACTCTCAATCTTGTGCTTGAAGTACTTCAAAACCCATCTCAAC 2880  
2821 TACATCCTGACTCTCAATCTTGTGCTTGAAGTACTTCAAAACCCATCTCAAC 2880  
2881 TCACCTGACTRTTTTACCCCAAGGTTTCAAGGATAGVCCCACTATTTGGCCAGGCAT 2940  
2881 TCACCTGACTRTTTTACCCCAAGGTTTCAAGGATAGVCCCACTATTTGGCCAGGCAT 2940  
2941 TAGCCCAAGACTTGAGYCARITYMTACATCCTGGACACTCTTGTCTCTTCRGTAAGTGATG 3000

2941 TAGCCCAAGACTTGAGYCARITYMTACATCCTGGACACTCTTGTCTCTTCRGTAAGTGATG 3000  
3001 ATTACTTTTTRCYGCVYRTTTCAGAAAACCTTGTGTCATCAAGCCACCCCAAGCRCTCTTMA 3060  
3001 ATTACTTTTTRCYGCVYRTTTCAGAAAACCTTGTGTCATCAAGCCACCCCAAGCRCTCTTMA 3060  
3061 ATTCTCTCGCYACCTGTGGCTACAWGTTTCCAAACAPARGCTCARCTCTGCTCAAGC 3120  
3061 ATTCTCTCGCYACCTGTGGCTACAWGTTTCCAAACAPARGCTCARCTCTGCTCAAGC 3120  
3121 AGTTTAAATACCTTAGBCTTAARATTTATCCAAAGGACCAAGGCCCTCAGTGGGAAYRYA 3180  
3121 AGTTTAAATACCTTAGBCTTAARATTTATCCAAAGGACCAAGGCCCTCAGTGGGAAYRYA 3180  
3181 TCCAGCTTATCTGCGCTTATCCTCATCYCAAAACCTTAAAGCAACTTAAAGRRRTTCTTTG 3240  
3181 TCCAGCTTATCTGCGCTTATCCTCATCYCAAAACCTTAAAGCAACTTAAAGRRRTTCTTTG 3240  
3241 GCTTAAYAGGYTTCTGCCGAATGATTTCCCAAGTWTGGCRAAATAGCCAGGYCATTA 3300  
3241 GCTTAAYAGGYTTCTGCCGAATGATTTCCCAAGTWTGGCRAAATAGCCAGGYCATTA 3300  
3301 WATACASTAATTAAGGAACTCAGAAAGCCATACCCATTTARTAGATGGAAMCTGAA 3360  
3301 WATACASTAATTAAGGAACTCAGAAAGCCATACCCATTTARTAGATGGAAMCTGAA 3360  
3361 GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
3361 GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
3421 TGCCAAACRGGGCAAGACTTTTSTTYATAYRTCAAGAAAACCAAGYCCCAGTGTTRGGA 3480  
3421 TGCCAAACRGGGCAAGACTTTTSTTYATAYRTCAAGAAAACCAAGYCCCAGTGTTRGGA 3480  
3481 GTCCCTTACACAGTCCBAGGAYGAGCTTTCGCAACCYRTGGCRVACCTGASTAGGAAAYT 3540  
3481 GTCCCTTACACAGTCCBAGGAYGAGCTTTCGCAACCYRTGGCRVACCTGASTAGGAAAYT 3540  
3541 GATGTAGTGGCAAGGTTGRCYTCATTTTAYGGGTAGTGTGGCAGTAGCAGTYKTA 3600  
3541 GATGTAGTGGCAAGGTTGRCYTCATTTTAYGGGTAGTGTGGCAGTAGCAGTYKTA 3600  
3601 GTATCTGAAGCAGTGTAAAAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 GTATCTGAAGCAGTGTAAAAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 GTGAAIRGCATCTACTCTCTAAGAGACTTGTGCTGTGTCAGACACACCTTACTTAAA 3720  
3661 GTGAAIRGCATCTACTCTCTAAGAGACTTGTGCTGTGTCAGACACACCTTACTTAAA 3720  
3721 TRTCAGGCTCTATTACTTGAAGGCCAGTGTGCTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
3721 TRTCAGGCTCTATTACTTGAAGGCCAGTGTGCTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
3781 CCAGYCNCACTTTCTTCAGACAATGAAGAAAAGATAPAAATATCTGTCACCAARTAAAT 3840  
3781 CCAGYCNCACTTTCTTCAGACAATGAAGAAAAGATAPAAATATCTGTCACCAARTAAAT 3840  
3841 TCTCAAACTTATGCGCTCGAGGGGACCTTCTTAGAGTTCCYTTGACTGATCCYGACCTT 3900  
3841 TCTCAAACTTATGCGCTCGAGGGGACCTTCTTAGAGTTCCYTTGACTGATCCYGACCTT 3900  
3901 CAACTTGATATCTAGTGAAGATTTCTTTGTAGAAAAGAGACTTCGAAAAGYGGGATGTC 3960  
3901 CAACTTGATATCTAGTGAAGATTTCTTTGTAGAAAAGAGACTTCGAAAAGYGGGATGTC 3960  
3961 AGTGGTCAGTGATAATGGAATATYTTGAAAGTAATCCCTCCTCAGTCCAGAACTAGTGTCTA 4020  
3961 AGTGGTCAGTGATAATGGAATATYTTGAAAGTAATCCCTCCTCAGTCCAGAACTAGTGTCTA 4020  
4021 GCTRGCAAGACTTAATAGCCCTCAATYKGGGCACTAGATTTAGGAGAAAGAAAGGGYAAA 4080  
4021 GCTRGCAAGACTTAATAGCCCTCAATYKGGGCACTAGATTTAGGAGAAAGAAAGGGYAAA 4080

Qy	4081	TATATATACAGACTCT	TRARTATGCTT	YACCTAGCT	TNTCATG	CCCATGMRGCAATATG	GSAR	4141					
Db	4081	TATATATACAGACTCT	TRARTATGCTT	YACCTAGCT	TNTCATG	CCCATGMRGCAATATG	GSAR	4140					
Qy	4141	AGAAAGGGAATTCCT	AACTTCY	GAGRGAAACCTT	ATCAMA	CATCAGGAAGCCATTAG	GAAR	4200					
Db	4141	AGAAAGGGAATTCCT	AACTTCY	GAGRGAAACCTT	ATCAMA	CATCAGGAAGCCATTAG	GAAR	4200					
Qy	4201	ATTATTAYTTGG	CGWGTACAGAA	CCCTAR	AGAGTG	GMAGTCTTCA	CTGCTGGGGTCA	TCA	4260				
Db	4201	ATTATTAYTTGG	CGWGTACAGAA	CCCTAR	AGAGTG	GMAGTCTTCA	CTGCTGGGGTCA	TCA	4260				
Qy	4261	NAAGGAAAGRAA	AGGGAATASA	AGRGAA	VTGCC	CAAGCAKATATT	GGAAGMAAAGAG	C	4320				
Db	4261	NAAGGAAAGRAA	AGGGAATASA	AGRGAA	VTGCC	CAAGCAKATATT	GGAAGMAAAGAG	C	4320				
Qy	4321	TGCAGG	CAGGACCCCT	CCATTAGAA	TGCTTATT	AAACTTCCCTTAGTAT	TAGGGTAA	TCC	4380				
Db	4321	TGCAGG	CAGGACCCCT	CCATTAGAA	TGCTTATT	AAACTTCCCTTAGTAT	TAGGGTAA	TCC	4380				
Qy	4381	CTTTCGGG	AAAAACCAAG	CCCCCAGT	ACTCAG	CAGGAGAAACAG	AGATG	GGGAACCTC	CAGGAG	4440			
Db	4381	CTTTCGGG	AAAAACCAAG	CCCCCAGT	ACTCAG	CAGGAGAAACAG	AGATG	GGGAACCTC	CAGGAG	4440			
Qy	4441	CAGTTT	TTCCCTCGG	ACGGTTAG	CCCACTG	GAAGAGG	AAAAATATCTTTG	CTCCTCAAC	4500				
Db	4441	CAGTTT	TTCCCTCGG	ACGGTTAG	CCCACTG	GAAGAGG	AAAAATATCTTTG	CTCCTCAAC	4500				
Qy	4501	TATCCA	ATGGAAAT	TACTTTAA	AAACCTTT	CATCAA	ACCTTTCACTTAG	GGCATCGATAG	CAC	4560			
Db	4501	TATCCA	ATGGAAAT	TACTTTAA	AAACCTTT	CATCAA	ACCTTTCACTTAG	GGCATCGATAG	CAC	4560			
Qy	4561	CCATCAR	ATGGC	CAAAATCAT	TATTAT	TCTGG	ACAGGCTTTTCA	AAACTAT	CAAGCAR	4620			
Db	4561	CCATCAR	ATGGC	CAAAATCAT	TATTAT	TCTGG	ACAGGCTTTTCA	AAACTAT	CAAGCAR	4620			
Qy	4621	AKTCAG	GGCCTGT	GAAATGT	GCCAR	AAAAATATCCCT	CTGCCTYAT	CGC	CAAGCTCCTTC	4680			
Db	4621	AKTCAG	GGCCTGT	GAAATGT	GCCAR	AAAAATATCCCT	CTGCCTYAT	CGC	CAAGCTCCTTC	4680			
Qy	4681	AGGAR	ACAAABA	CAGGCCAT	TACCT	GRABA	ARACTG	GC	CAACTGATTTT	CCCCACAAG	4740		
Db	4681	AGGAR	ACAAABA	CAGGCCAT	TACCT	GRABA	ARACTG	GC	CAACTGATTTT	CCCCACAAG	4740		
Qy	4741	CCCAAA	CCCTCAG	GGATTTCAG	TATCT	ACTAGTCT	CGGTAR	ATACTTT	CACGGGTT	GGGCA	4800		
Db	4741	CCCAAA	CCCTCAG	GGATTTCAG	TATCT	ACTAGTCT	CGGTAR	ATACTTT	CACGGGTT	GGGCA	4800		
Qy	4801	PAG	GCCTTCCC	TGTAG	GACAGAA	AGGCCCAAG	AGGTAA	TAAAG	GCCTAGTTC	ATGAA	4860		
Db	4801	PAG	GCCTTCCC	TGTAG	GACAGAA	AGGCCCAAG	AGGTAA	TAAAG	GCCTAGTTC	ATGAA	4860		
Qy	4861	ATA	ATTCC	CAGATTC	CGA	CTTCCCG	AGGCTTAC	AGAGTGA	CAATAG	CCCTGCTTTCCAG	4920		
Db	4861	ATA	ATTCC	CAGATTC	CGA	CTTCCCG	AGGCTTAC	AGAGTGA	CAATAG	CCCTGCTTTCCAG	4920		
Qy	4921	GCC	CAGTAA	CCCGAG	GATATCC	CGGCTTAG	GTATAC	GATATCA	CTTAC	TGCGCC	4980		
Db	4921	GCC	CAGTAA	CCCGAG	GATATCC	CGGCTTAG	GTATAC	GATATCA	CTTAC	TGCGCC	4980		
Qy	4981	TGA	AGGCC	CACAGT	CTCAG	GGAAGGT	CGAG	MAATGA	ATGA	AACTCAA	AGGCATCTA	5040	
Db	4981	TGA	AGGCC	CACAGT	CTCAG	GGAAGGT	CGAG	MAATGA	ATGA	AACTCAA	AGGCATCTA	5040	
Qy	5041	AAAA	AGCAAA	CCCGAG	GAACCC	CACTC	ACATG	CCCTG	YTTCTG	TTCCTAT	TAGCCTT	AAAA	5100
Db	5041	AAAA	AGCAAA	CCCGAG	GAACCC	CACTC	ACATG	CCCTG	YTTCTG	TTCCTAT	TAGCCTT	AAAA	5100
Qy	5101	AGA	ATCTG	CAACTTTCC	CAAA	AGCAG	CACTTAG	CCCAT	AGAA	TGCTGT	TGGA	AGG	5160
Db	5101	AGA	ATCTG	CAACTTTCC	CAAA	AGCAG	CACTTAG	CCCAT	AGAA	TGCTGT	TGGA	AGG	5160

Qy	5161	CCCTTCATACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGTCAGACATCA	5220
Db	5161	CCCTTCATAACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGTCAGACATCA	5220
Qy	5221	CCTCCTTAGCCAAATATCAAAAGTCTTAAAAAATTACAAGGAACCTATCCCTGAGAAG	5280
Db	5221	CCTCCTTAGCCAAATATCAAAAGTCTTAAAAAATTACAAGGAACCTATCCCTGAGAAG	5280
Qy	5281	AGGAAAGAACTATTCCACCCWWTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT	5340
Db	5281	AGGAAAGAACTATTCCACCCWWTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT	5340
Qy	5341	CCCCATCCTAGATACATCTCTGGGAAGAACCTTATTTATTTATACCCCAACTG	5400
Db	5341	CCCCATCCTAGATACATCTCTGGGAAGAACCTTATTTATTTATACCCCAACTG	5400
Qy	5401	CGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTGAGTCAAAATCTGGATAC	5460
Db	5401	CGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTGAGTCAAAATCTGGATAC	5460
Qy	5461	TGCCAAAGGAACCTGAAAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACTCTAGAGG	5520
Db	5461	TGCCAAAGGAACCTGAAAAATCCAGGAGACAACGCTAGCTATTCTCTGTGAACTCTAGAGG	5520
Qy	5521	ATTGTGGCTGCTCTTCAAAACAACAACAGGAGGAAAGTAACATAAAATCATAAATCCCCC	5580
Db	5521	ATTGTGGCTGCTCTTCAAAACAACAACAGGAGGAAAGTAACATAAAATCATAAATCCCCC	5580
Qy	5581	ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTGTTSTTTTACCTSTTTCACCTCTCACT	5640
Db	5581	ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTGTTSTTTTACCTSTTTCACCTCTCACT	5640
Qy	5641	GCACCCCTCCATGCGCGCTGTATGACCAAGTAGTCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Db	5641	GCACCCCTCCATGCGCGCTGTATGACCAAGTAGTCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Qy	5701	ATGCAAGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCCC	5760
Db	5701	ATGCAAGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAAACCCCC	5760
Qy	5761	ACCTTCACCTGCCACACCCATATGCCCGCCTACTATACATCTGCACTCTTTTGCATG	5820
Db	5761	ACCTTCACCTGCCACACCCATATGCCCGCCTACTATACATCTGCACTCTTTTGCATG	5820
Qy	5821	CATGCAAAATPACTCATATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGGACTT	5880
Db	5821	CATGCAAAATPACTCATATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGGACTT	5880
Qy	5881	GGAGTCACTGCTGTGTGGACTTACTTCAACCACTGGTATGCTCATGCGGGTGGAGTT	5940
Db	5881	GGAGTCACTGCTGTGTGGACTTACTTCAACCACTGGTATGCTCATGCGGGTGGAGTT	5940
Qy	5941	CAAGATCAGGCAAGGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCSGGGTACAT	6000
Db	5941	CAAGATCAGGCAAGGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCSGGGTACAT	6000
Qy	6001	GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAATCATCAAAACCCCTCCGTACC	6060
Db	6001	GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAATCATCAAAACCCCTCCGTACC	6060
Qy	6061	CATACTGCGCTGGTAAGCCTATTTAATACCAACCTCCTGAGGGCTCCATGAGGTCTCGGCC	6120
Db	6061	CATACTGCGCTGGTAAGCCTATTTAATACCAACCTCCTGAGGGCTCCATGAGGTCTCGGCC	6120
Qy	6121	CAAAACCTTACTACTGTGTGGATATGCTTCCCTCGAATCTCARGCCATATGTTTCAATC	6180
Db	6121	CAAAACCTTACTACTGTGTGGATATGCTTCCCTCGAATCTCARGCCATATGTTTCAATC	6180
Qy	6181	CCTGTACTGTAAACAAATGGAACAACCTTCAGCAGCAGAAATAAACCACCTTCCGTTTTAGTA	6240
Db	6181	CCTGTACTGTAAACAAATGGAACAACCTTCAGCAGCAGAAATAAACCACCTTCCGTTTTAGTA	6240
Qy	6241	GGACCTCTGTGTTCCAAATSTGGAATAAACCCATACCTCAAAACCTCCTGTTGTAATAATT	6300

6241	Db		GGACCTCTTGTTCCCAATSTGGAATAAACCCATACCTCAACACCTCACTGTGTAAATTT	6300
6301	Qy		AGCAATACTACATACACAACAACCTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA	6360
6301	Db		AGCAATACTACATACACAACAACCTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA	6360
6361	Qy		ATAGTCTGCCTTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTGTGTTG	6420
6361	Db		ATAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTGTGTTG	6420
6421	Qy		AATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTATAGTGCCCCYATGRCATCTAC	6480
6421	Db		AATGGCTCTTCAGAAATCTATGTGCTTCTCTCAATCTTATAGTGCCCCYATGRCATCTAC	6480
6481	Qy		ACTGAAACAAGATTTATACAGTTATGTCATATCTAAAGCCCGCAACAAAAGAGTACCATT	6540
6481	Db		ACTGAAACAAGATTTATACAGTTATGTCATATCTAAAGCCCGCAACAAAAGAGTACCATT	6540
6541	Qy		CTTCTTTTTGTATAGGACGAGGAGTGCTAGGTGCATAGGTACTGGCATTTGGCGGTATC	6600
6541	Db		CTTCTTTTTGTATAGGACGAGGAGTGCTAGGTGCATAGGTACTGGCATTTGGCGGTATC	6600
6601	Qy		ACAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGG	6660
6601	Db		ACAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGG	6660
6661	Qy		GTCCGCGACTCCCTGGTTCACCTTGCAAGATCAACTTAACTCCCTGACGACAGTAGTCCCT	6720
6661	Db		GTCCGCGACTCCCTGGTTCACCTTGCAAGATCAACTTAACTCCCTGACGACAGTAGTCCCT	6720
6721	Qy		CRAAATCGAAGAGCTTTTAGACTGYCTAAACCGCTGARAGAGGGGAACTCTTTATTTTAA	6780
6721	Db		CRAAATCGAAGAGCTTTTAGACTGYCTAAACCGCTGARAGAGGGGAACTCTTTATTTTAA	6780
6781	Qy		GGGGAAGAATCCTGTTATATGTTTAAATCAATTCGGAAATCGTCACTGAGAAGTTTAAAGAA	6840
6781	Db		GGGGAAGAATCCTGTTATATGTTTAAATCAATTCGGAAATCGTCACTGAGAAGTTTAAAGAA	6840
6841	Qy		ATTCGAGATCGAATACAACTGAKAGCAGARGAGCTTCGAAACACTGAGACCCCTGGGGCTC	6900
6841	Db		ATTCGAGATCGAATACAACTGAKAGCAGARGAGCTTCGAAACACTGAGACCCCTGGGGCTC	6900
6901	Qy		CTCAGCCCATGGATGGCCCTGGATTTCTCCCTCTTTTAGGACCTCTAGCAGCTATAATATTG	6960
6901	Db		CTCAGCCCATGGATGGCCCTGGATTTCTCCCTCTTTTAGGACCTCTAGCAGCTATAATATTG	6960
6961	Qy		CTACTCTCTTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTTGTCTCTTCCAGAAATC	7020
6961	Db		CTACTCTCTTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTTGTCTCTTCCAGAAATC	7020
7021	Qy		GAAGCTGTAAACTTACAAATGGAGCCCAAGATGCAAGTCCAGACTAGACTATACCCGAGA	7080
7021	Db		GAAGCTGTAAACTTACAAATGGAGCCCAAGATGCAAGTCCAGACTAGACTATACCCGAGA	7080
7081	Qy		CCCCTGGACCGCGCTGYTAGCCCAAGATCTCATGTTAATGACATCAAGGACCCCTCTCT	7140
7081	Db		CCCCTGGACCGCGCTGYTAGCCCAAGATCTCATGTTAATGACATCAAGGACCCCTCTCT	7140
7141	Qy		GAGGAAATCTCAGCTGCAACAACCTCTACTAGCCCCCAATTCAGCAGGAAGCAGTTAGAGC	7200
7141	Db		GAGGAAATCTCAGCTGCAACAACCTCTACTAGCCCCCAATTCAGCAGGAAGCAGTTAGAGC	7200
7201	Qy		GGTGTGCGCCAAACCTCCCAACAGCACTTAGTTTTTCTCTTTGAGATGGGGGACTGAGA	7260
7201	Db		GGTGTGCGCCAAACCTCCCAACAGCACTTAGTTTTTCTCTTTGAGATGGGGGACTGAGA	7260
7261	Qy		GACAGGATAGCTGATTTCTTAGGCTCAYTAAGAAATCCYTAAGCTAGSTGGGAAGGTG	7320
7261	Db		GACAGGATAGCTGATTTCTTAGGCTCAYTAAGAAATCCYTAAGCTAGSTGGGAAGGTG	7320
7321	Qy		ACCACATCCACTTTTAAACACGGGGCTTGCAACTTAGTGTACACTGACCAATCAGAGAG	7380

Db	7321	ACACATCCACCTTTAAACACGGGGCTTGCACTT	AGTTCACACCTGACCAATCAGAGAG	7338	
Qy	7381	CTCACTAAAATGCTAATTAGGCAAGACAGAGG	ATAGCAAAATCAATTTATTGC	7440	
Db	7381	CTCACTAAAATGCTAATTAGGCAAGACAGAGG	ATAGCAAAATCAATTTATTGC	7440	
Qy	7441	MTGAGAGCAGCAGGAGGGACAATGATCGG	GATATAAACCCCAAGTTCGAGCCGGCAA	7500	
Db	7441	MTGAGAGCAGCAGGAGGGACAATGATCGG	GATATAAACCCCAAGTTCGAGCCGGCAA	7500	
Qy	7501	CGGCAACCCCTTTGGTCCCTCCCTTCCTTTG	TATGCGAGCTCTGTTTCATGCTATTTCAC	7560	
Db	7501	CGGCAACCCCTTTGGTCCCTCCCTTCCTTTG	TATGCGAGCTCTGTTTCATGCTATTTCAC	7560	
Qy	7561	TCTATTAAATCTTGCARCTGCR	7582		
Db	7561	TCTATTAAATCTTGCARCTGCR	7582		
RESULT 5	AY101582	10222 bp	DNA	linear	PRI 11-FEB-2004
LOCUS	AY101582	Homo sapiens isolate 21	endogenous retrovirus HERV-W, ERVWE1 locus,		
DEFINITION	allele A, complete sequence.				
ACCESSION	AY101582				
VERSION	AY101582.1	GI:37544399			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 10222)				
TITLE	Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.				
JOURNAL	The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology				
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)				
REFERENCE	14757826				
AUTHORS	1 (bases 1 to 10222)				
TITLE	Mallet, F., Bouton, O. and Oriol, G.				
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France				
FEATURES	Location/Qualifiers				
source	1..10222				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/isolate="21"				
	/isolation_source="PBMC of individual 21"				
	/db_xref="taxon:9606"				
	/chromosome="7"				
	/map="7q21-q22"				
	/sex="female"				
	/cell_type="PBMC"				
	/note="isolated by PCR; ERVWE1 locus, allele A endogenous_virus: HERV-W"				
LTR	1..780				
mRNA	join(248..811,2887..9813)				
	/note="putative mRNA transcript 1; gag-pol pseudogene"				
mRNA	join(248..811,7585..9813)				
	/product="envelope glycoprotein"				
	/note="putative mRNA transcript 2"				
mRNA	join(248..811,9248..9813)				
	/note="putative mRNA transcript 3; no coding region"				
CDS	7812..9428				
	/note="syncytin"				
	/codon_start=1				
	/product="envelope glycoprotein"				
	/protein_id="AA068161.1"				
	/db_xref="GI:37544400"				
	/translation="MALPYHFLFTVLLPSTLTAPPCCRCWTSSPYOETLWMQRQI GNIDAPYSRSLSKGTPTTATHTMPRNCIHSATLCMHANTHWGKMINSPCGGLG				



LTR  
ORIGIN

Qy	606	GATGGGAAACGTTCCCGCAGAACAAAAACGCCCTTAAGACGTATTCTGGARAAATGGGA	665
Db	2886	GATGGGAAACGTTCCCGCAGAACAAAAACGCCCTTAAGACGTATTCTGGARAAATGGGA	2945
Qy	666	MCATTTTGACCCCTCAGACACATTAAGAAAGAACGACTTATATTCTTCTGCACTGCCGCTG	725
Db	2946	CCAATTTGACCCCTCAGACACATTAAGAAAGAACGACTTATATTCTTCTGCACTGCCGCTG	3005
Qy	726	GCATCTCTGAGGGAAGTATAAAATTATAAACACCATCTTACAGCTAGACYCTCTTTGTAGAA	785
Db	3006	GCATCTCTGAGGGAAGTATAAAATTATAAACACCATCTTACAGCTAGACCTCTTTGTAGAA	3065
Qy	786	AAGCAAAATGGAGTGAAGTGCCTAAGTAAGTACAAACTTTCTTTTCAATTAAGAGACAACTCAC	845
Db	3066	AAGCAAAATGGAGTGAAGTGCCTAAGTAAGTACAAACTTTCTTTTCAATTAAGAGACAACTCAC	3125
Qy	846	AAATTATCTAAAAAGTGTGATTATATGCCCTACAGGAGCCCTTCAGAGTCTACCTCCCTATC	905
Db	3126	AAATTATCTAAAAAGTGTGATTATATGCCCTACAGGAGCCCTTCAGAGTCTACCTCCCTATC	3185
Qy	906	CCAGCATCCCGCACTCTCTCCCAAAATAAAGGACCCCTCTTCAACCCCAATAGTCCAA	965
Db	3186	CCAGCATCCCGCACTCTCTCCCAAAATAAAGGACCCCTCTTCAACCCCAATAGTCCAA	3245
Qy	966	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATTCCTCCCAATTATGA	1025
Db	3246	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCAAATATTCCTCCCAATTATGA	3305
Qy	1026	CCCTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCACGACGAGTGCTGTCGCTTTTYY	1085
Db	3306	CCCTC-CCAAGCAGTGGGAGGAGAGAAATTCGGCCACGACGAGTGCTGTCGCTTTTTC	3364
Qy	1086	TCTCCCAAGCTTAAAGCAAAATAAAAAACAGACTTAGGTAAATTCCTCAGATAAYCCTGATGG	1145
Db	3365	TCTCCCAAGCTTAAAGCAAAATAAAAAACAGACTTAGGTAAATTCCTCAGATAACCTGATGG	3424
Qy	1146	CTATATTGRTGTTTTTACAAGGGTTAGACAATTCCTTTGATCTGACATGGAGAGATATATA	1205
Db	3425	CTATATTGATGTTTTTACAAGGGTTAGACAATTCCTTTGATCTGACATGGAGAGATATA-A	3483
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCATAACTGCAGCCT	1265
Db	3484	TGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCATAACTGCAGCCT	3543
Qy	1266	GAGRGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGGATGCAACAGAAAGG	1325
Db	3544	GAGAGTTTGGCGATCTCTGTATCTCTCAGTCAGGTCAATG-AATAGGATGACAACAG-AGG	3600
Qy	1326	AAAGANNAATGATTTCCCAACAGGCCAGCAGCAGTTCCTCAGTCTASACCTCATTTGGGAC	1385
Db	3601	AAAGAGAAATGATTTCCCAACAGGCCAGCAGCAGTTCCTCAGTCTAGACCTCATTTGGG---	3657
Qy	1386	ACAGAAATCAGTAAACATGGGAGAGATTGGTGTGCGAGACATTTGCTTAACCTGTGTGCTASAA	1445
Db	3658	ACACAGAAACAGAAACAT-GGAGATTTGGTGTGTGCGAGACATTTGCTTAACCTGTGTGCTAGAA	3716
Qy	1446	GGACTAAGGAAAACTASGAAAGAAATCTAYGAATTAATCTCAATGATGTCCACCAATAACACA	1505



Db	6976	AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTAGATATCTTCA	CGGGTTGGCGCAGAG	7035
QY	4804	GCCTTCCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCAC	TAGTTTCATGAATA	4863
Db	7036	GCCTTCCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCAC	TAGTTTCATGAATA	7095
QY	4864	ATTCCACAGATTCGGACTCCCGCAGGCTTACAGAGTGA	CAATAGCCCTGCTTCCAGGCC	4923
Db	7096	ATTCCACAGATTCGGACTCCCGCAGGCTTACAGAGTGA	CAATAGCCCTGCTTCCAGGCC	7155
QY	4924	ACAGTAAACCAGGGAGTATCCCAGCGTTAGGTATACGATATCACTT	ACACTGCGCCTGA	4983
Db	7156	ACAGTAAACCAGGGAGTATCCCAGCGTTAGGTATACGATATCACTT	ACACTGCGCCTGA	7215
QY	4984	AGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGAATGAAATYAC	TCAAAGGACATCTAAAA	5043
Db	7216	AGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGAATGAAATYAC	TCAAAGGACATCTAAAA	7275
QY	5044	AAGCAAAACCCAGGAACCCACCTCACATGGCCTGYCTGTGTGCT	TATAGCCTTAAAAAGA	5103
Db	7276	AAGCAAAACCCAGGAACCCACCTCACATGGCCTGYCTGTGTGCT	TATAGCCTTAAAAAGA	7335
QY	5104	ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATG	CTGATGCAAGGCC	5163
Db	7336	ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATG	CTGATGCAAGGCC	7395
QY	5164	TTCTATAACCAATGACTTGTGCTTTGACCCACAGACAGCCAACTT	AGTTTGCAGACATCACCT	5223
Db	7396	TTCTATAACCAATGACTTGTGCTTTGACCCACAGACAGCCAACTT	AGTTTGCAGACATCACCT	7455
QY	5224	CTTTAGCCCAATATCAACAGTTCTTTAAACAATTAAGGAACTTAT	TCCTGAGAGAGG	5283
Db	7456	CTTTAGCCCAATATCAACAGTTCTTTAAACAATTAAGGAACTTAT	TCCTGAGAGAGG	7515
QY	5284	GAAGAAGAACTATTCCACCCMWGTGACATGTTATTAGTCAAGTC	CCCTCYCTCTAAATCCC	5343
Db	7516	GAAGAAGAACTATTCCACCCMWGTGACATGTTATTAGTCAAGTC	CCCTCYCTCTAAATCCC	7575
QY	5344	CATCCCTAGATATCATCTCTGGGAAGGACCCCTACCCAGTCA	TTTATYATACCCCAACTGCGG	5403
Db	7576	CATCCCTAGATATCATCTCTGGGAAGGACCCCTACCCAGTCA	TTTATYATACCCCAACTGCGG	7635
QY	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATATATCACACTTGAGTCA	AAATCTGTGATCTG	5463
Db	7636	TTAAAGTGGCTGGAGTGGAGTCTTGGATATATCACACTTGAGTCA	AAATCTGTGATCTG	7695
QY	5464	CAAAGGAACCTGAAATCTCAGGAGACAAACCTAGCTATTCTCTG	TGAACCTCTAGAGATT	5523
Db	7696	CAAAGGAACCTGAAATCTCAGGAGACAAACCTAGCTATTCTCTG	TGAACCTCTAGAGATT	7755
QY	5524	TGGCCTGCTCTTCAAAACAACCAAGGAGGAAGTAACATAAATCA	TAAATCCCCCATTG	5583
Db	7756	TGGCCTGCTCTTCAAAACAACCAAGGAGGAAGTAACATAAATCA	TAAATCCCCCATTG	7814
QY	5584	GSCTCCCTTATCATATTTTCTCTKTATGTTTACCSTTTTCACTC	TCATCTCATGCA	5643
Db	7815	GSCTCCCTTATCATATTTTCTCTKTATGTTTACCSTTTTCACTC	TCATCTCATGCA	7874
QY	5644	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACGAGAGTTT	CTATGGAGATG	5703
Db	7875	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACGAGAGTTT	CTATGGAGATG	7934
QY	5704	CAGGCTCCCGGAATATTTGATGCCCATCGTATAGAGTCTTTSTA	AGGGAACCCCCACC	5763
Db	7935	CAGGCTCCCGGAATATTTGATGCCCATCGTATAGAGTCTTTSTA	AGGGAACCCCCACC	7994
QY	5764	TTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACT	CTTTGCAATGAT	5823
Db	7995	TTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACT	CTTTGCAATGAT	8054
QY	5824	GCAAACTACTATTATGGACAGAAAAATGAATTAATCCTTAGTTG	TCTCGAGGACTTGA	5883

8055	QY	GTCACTGCTGTTGGACCTTACTTGGACAGGAAAAAATGATTAATCTCTAGTTGCTCGGAGGACCTTGGCA	8114
5984	QY	GTCACTGCTGTTGGACCTTACTTGGACAGGAAAAAATGATTAATCTCTAGTTGCTCGGAGGACCTTGGCA	5943
8115	QY	GTCACTGCTGTTGGACCTTACTTGGACAGGAAAAAATGATTAATCTCTAGTTGCTCGGAGGACCTTGGCA	8174
5944	QY	GATCAGGCAAGAGAGAAAAACATGTAAGAGAAAGTAATCTCCAACTCACCGGGGTACATGGC	6003
8175	QY	GATCAGGCAAGAGAGAAAAACATGTAAGAGAAAGTAATCTCCAACTCACCGGGGTACATGGC	8234
6004	QY	ACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	6063
8235	QY	ACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	8294
6064	QY	ACTCGCCTGTGTAGCCTATTATTAACACACCTCACTGGGCTCCATGAGGTCTCGGCCCAA	6123
8295	QY	ACTCGCCTGTGTAGCCTATTATTAACACACCTCACTGGGCTCCATGAGGTCTCGGCCCAA	8354
6124	QY	AACCCCTACTAATCTGTTGGATATGCTCCCTCGTGAACCTTCARGGCCATATGTTTCAATCCCT	6183
8355	QY	AACCCCTACTAATCTGTTGGATATGCTCCCTCGTGAACCTTCARGGCCATATGTTTCAATCCCT	8414
6184	QY	GTACCTGGAACAATGGAAACAATCTCAGCAGAGAAATAAACACACCTTCGGTTTTAGTAGGA	6243
8415	QY	GTACCTGGAACAATGGAAACAATCTCAGCAGAGAAATAAACACACCTTCGGTTTTAGTAGGA	8474
6244	QY	CCTCTGTTTCCAAATSTGGAAAAATAACCCATACCTCAACACCTCACTGTGTAAAAATTTAGC	6303
8475	QY	CCTCTGTTTCCAAATSTGGAAAAATAACCCATACCTCAACACCTCACTGTGTAAAAATTTAGC	8534
6304	QY	AATPACTACATACAACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCCAACACAAATA	6363
8535	QY	AATPACTACATACAACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCCAACACAAATA	8594
6364	QY	GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGAAT	6423
8595	QY	GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGAAT	8654
6424	QY	GGCTCTTCAGAAATCTATGTCTCTCTCATCTTAGTGCCCCVATGRCCATCTACACT	6483
8655	QY	GGCTCTTCAGAAATCTATGTCTCTCTCATCTTAGTGCCCCVATGRCCATCTACACT	8714
6484	QY	GAAAGAAATTTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCATCTCT	6543
8715	QY	GAAAGAAATTTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCATCTCT	8774
6544	QY	CTTTTGTATTAGGACAGGAGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATACACA	6603
8775	QY	CTTTTGTATTAGGACAGGAGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATACACA	8834
6604	QY	ACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGGACATGGACGGGTC	6663
8835	QY	ACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGGACATGGACGGGTC	8894
6664	QY	GCCGATCCCTGGTCACTTGCAGAGTCAACTTTAACTCCCTAGCAGCAGTAGTCTCTTCA	6723
8895	QY	GCCGATCCCTGGTCACTTGCAGAGTCAACTTTAACTCCCTAGCAGCAGTAGTCTCTTCA	8954
6724	QY	AATCGAAGAGCTTTAGACTTGTCTTACCGCTGABAGAGGGGGAACCTGTTTATTTTAGGG	6783
8955	QY	AATCGAAGAGCTTTAGACTTGTCTTACCGCTGABAGAGGGGGAACCTGTTTATTTTAGGG	9014
6784	QY	GAAGAAATGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTCAGAGAAAGTTTAAAGAAAT	6843
9015	QY	GAAGAAATGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTCAGAGAAAGTTTAAAGAAAT	9074
6844	QY	CSAGATCGAAATACACGTAACAGCAGAGAGCTTCGAAAAACCTGGACCCCTGGGGCCCTCCTC	6903
9075	QY	CSAGATCGAAATACACGTAACAGCAGAGAGCTTCGAAAAACCTGGACCCCTGGGGCCCTCCTC	9134
6904	QY	AGCCBAATGGATGCCCTGGATCTCCCTCTCTTAGGACCTCTAGCAGGTATTAATTTGCTA	6963
9135	QY	AGCCBAATGGATGCCCTGGATCTCCCTCTCTTAGGACCTCTAGCAGGTATTAATTTGCTA	9194

```
6964 CTCTCTTTGGACCTCTATCTTTTACCTCTCTGTTAACTTTGTCTCTTCCAGAAATCGAA 7023
Db CTTCTCTTTGGACCTCTATCTTTTAACTCTCTGTTAACTTTGTCTCTTCCAGAAATCGAA 9254
Qy GCTGTTAAACTACAAATGGAGCCCAAGATGAGTCCCAAGACTAAGATCTACCGCAGACCC 7083
Db GCTGTTAAACTACAAATGGAGCCCAAGATGAGTCCCAAGACTAAGATCTACCGCAGACCC 9314
Qy CTGACCGGCTGTAGTCCACAGCTCTGATGTTAATGACATCAATCAAGAGGACCCCTCTCTGAG 7143
Db CTGACCGGCTGTAGTCCACAGCTCTGATGTTAATGACATCAATCAAGAGGACCCCTCTCTGAG 9374
Qy GAAATCTCAGCTGACCAACCTCTACTAGCCGCTTACGAGGAGGAGGAGGAGGAGGAGGAG 7203
Db GAAATCTCAGCTGACCAACCTCTACTAGCCGCTTACGAGGAGGAGGAGGAGGAGGAGGAG 9434
Qy SGTGCGCAACCTCCCAACAGCACTTACTAGTCTTCTGTTGAGATGGGGGAGTCTGAGAGAC 7263
Db SGTGCGCAACCTCCCAACAGCACTTACTAGTCTTCTGTTGAGATGGGGGAGTCTGAGAGAC 9494
Qy AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCVTAAGCCTAGTGGGAGGAGTCAAC 7323
Db AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCVTAAGCCTAGTGGGAGGAGTCAAC 9554
Qy AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCVTAAGCCTAGTGGGAGGAGTCAAC 7383
Db AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCVTAAGCCTAGTGGGAGGAGTCAAC 9614
Qy ACTAAATGCTAATAGGCAAGACAGGAGTAAAGAAATAGCAATCAATCAATCAATCAATCA 7443
Db ACTAAATGCTAATAGGCAAGACAGGAGTAAAGAAATAGCAATCAATCAATCAATCAATCA 9674
Qy AGACACAGCAGGAGGAGCAATGATCGGGATATAAACCAGAGTTCGAGCGGCAACGG 7503
Db AGACACAGCAGGAGGAGCAATGATCGGGATATAAACCAGAGTTCGAGCGGCAACGG 9734
Qy CAACCCCTTTGGGTCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db CAACCCCTTTGGGTCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACTCT 9794
Qy ATTAAATCTGCARTGCR 7582
Db ATTAAATCTGCARTGCA 9813

RESULT 6
AY101585 10222 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVWE1 locus,
DEFINITION allele B, complete sequence.
ACCESSION AY101585.1 GI:37544405
VERSION 1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10222)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 10222)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10222
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="79"
/isolation_source="PBMC of individual 79"
/db_xref="taxon:9606"
/chronosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC"
/notes="caucasian
isolated by PCR; ERVWE1 locus, allele B
endogenous_virus: HERV-W"
1..780
join(248..811,2887..9813)
/notes="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/notes="putative mRNA transcript 2"
/notes="putative mRNA transcript 2"
join(248..811,9248..9813)
/notes="putative mRNA transcript 3; no coding region"
7812..9428
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AM68164.1"
/db_xref="GI:37544406"
/translating="MALPHYFLFTVLLPFTLTAPPKCMTSSSPYQBLWRMQR
GNIDAPSYSLSGTPTTAHTMPNCHVSATLCMHANTHYTMKMINPSCGGLG
TVCTWYFTQMSDGGVDOAREKHKVVISOLTRVHGTSSPYKGLDLSKLEHLLRT
HTLVSLPNTTLGLHEVSAONPCNKVICLPLNFRPVSI PVPEOWNFSTEINTTSV
LVGPLVNLBITHTSHLTCVKFNTYTTNSQIRWTPPTQIVCLPSGIFVCGTSA
YRCINGSSSMCFSLFPPMTIYTDQDLYSVISPRNKRPVLPILFVIGAGVJGALG
TGCTGTLPLGECYVNGSIGIVTEKVEIRDIQRAEELRNTGPMGLLSQWMPWILP
RGCTPLAALILLPLPCIFNLVNFVSSRIEAVKLQMEPKMQSKTIYRRLDPRPAS
RSDVNDIKGTPPEISNAQPLLRNSAGS"
9487..10222
```

LTR  
ORIGIN

```
Query Match 85.0%; Score 6448; DB 9; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

Qy 606 GATGGGAAAGCTTCCCGCAGACAAAACGCCCCTAAGACGCTTCTTGGARAATTGGGA 665
Db 2886 GATGGGAAAGCTTCCCGCAGACAAAACGCCCCTAAGACGCTTCTTGGARAATTGGGA 2945

Qy 666 MCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTTCTGACGTGCGCCTG 725
Db 2946 CCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTTCTGACGTGCGCCTG 3005

Qy 726 GCCTCTCTGAGGAGTAAATTAATTAACCATCTTACAGCTAGACVCTTTTGTAGAA 785
Db 3006 GCCTCTCTGAGGAGTAAATTAATTAACCATCTTACAGCTAGACVCTTTTGTAGAA 3065

Qy 786 AAGCAAATGAGTGAAGTGCCATAAGTACAACTTTTTCATTAAGACAACTCAC 845
Db 3066 AAGCAAATGAGTGAAGTGCCATAAGTACAACTTTTTCATTAAGACAACTCAC 3125

Qy 846 AATTATGTAAAGAGTGATTTATGCTCCTACAGGAAGCCTTCAGAGTCTACTCCCTATC 905
Db 3126 AATTATGTAAAGAGTGATTTATGCTCCTACAGGAAGCCTTCAGAGTCTACTCCCTATC 3185

Qy 906 CCAGCATCCCCGACTCTCTTCCCCAMTAAATAGGACCCCTTCAACCCAAATGGTCCAA 965
Db 3186 CCAGCATCCCCGACTCTCTTCCCCAACTTAATAGGACCCCTTCAACCCAAATGGTCCAA 3245

Qy 966 AAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCAATATTTCCCAATATGA 1025
Db 3246 AAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCAATATTTCCCAATATGA 3305

Qy 1026 CCCCCTCCCAACGAGTGGGAGGAGGAATTCGGCCCAAGAGTGCATGTCYTTTYY 1085
Db 1026 CCCCCTCCCAACGAGTGGGAGGAGGAATTCGGCCCAAGAGTGCATGTCYTTTYY 1085
```

3306	CCCCT-CCAAAGCAGTGGGAGGAAGAAATTCGGCCACCCAGAGTGCATGTGCTTTTTC	3364
1086	TCCTCCAGACTTAAAGCAAAATAAAAAAGACACTTAGETAAATTTCTCAGATAAYCCTGATGG	1145
3365	TCCTCCAGACTTAAAGCAAAATAAAAAAGACACTTAGETAAATTTCTCAGATAACCCCTGATGG	3424
1146	CTATATTGRTGTTTTACAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAGATATATA	1205
3425	CTATATTGATGTTTTACAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAGATATA-A	3483
1206	TGTCATCTCTAAATCAGACACTTAACCCCAATGAGAGAGTGCACACATTAATCTGCAGCCT	1265
3484	TGTCATCTCTAAATCAGACACTTAACCCCAATGAGAGAGTGCACACATTAATCTGCAGCCT	3543
1266	GAGRGTTTTGGCGATCTCTGGTATCTCAGTCAGETCAATGGATANGATGACACAAGAGG	1325
3544	GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGETCAATG--ATAGGATGACAACAG-AGG	3600
1326	AAAGNNATGATTTCCCAACAGGCCAGCARGCAGTTCCTCAGCTCTASACCCCTCATTTGGGAC	1385
3601	AAAGAGATGATTTCCCAACAGGCCAGCARGCAGTTCCTCAGCTCTAGACCCCTCATTTGG--	3657
1386	ACAGAAATCAGTAAATCGGAGATGGTCTGCAGACATTTGCTAACTTGTGTGCTASAA	1445
3658	ACACAGAAATCAACAACAT-GGAGATTGGTGTGCACACATTTGCTAACTTGTGTGCTAGAA	3716
1446	GGACTAAGGAAAACTASGAAGAAATCTATGAATTTACTCAATGATGTCACACCATACACA	1505
3717	GGACTAAGAAATCTAGGAG-AGTCTATGAATTTCTCAATGATGTCACCATACACA	3775
1506	GGGAAGGGGAAGAAAATCTCTACTGCCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC	1565
3776	-GGGAAGGGGAAGAAAATCTCTACTGCCCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC	3834
1566	GTGCTCTCTGTGCACCTGCATCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTTATCA	1625
3835	GTGCTCTCTGTGCACCTGCATCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTTATCA	3894
1626	CTCAGTCAGCTGCAGACATTAG-AAAAAACTTCAAAAGTCTGCGGTAGGCCCGGAGCAAA	1684
3895	CTCAGTCAGCTGCAGACATTAGAAAAAAACTTTCAAAGTCTGCGGTAGGCCCGGAGCAAA	3954
1685	ACTTAGAAAACCTATTGAACTTGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC	1744
3955	ACTTAGAAAACCTATTGAACTTGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC	4014
1745	AGCGGAACAGGACAAACGGGATTAATAAAAAAGCCACCGCTTTAGTCATGACCCCTCAGG	1804
4015	AGCGGAACAGGACAAACGGGATTAATAAAAAAGCCACCGCTTTAGTCATGACCCCTCAGG	4074
1805	CAAGTGACATTTGGAGGCTCTGGAAGAGGAAAAGCTGGGCAATTTGAATGCTTAATAGG	1864
4075	CAAGTGACATTTGGAGGCTCTGGAAGAGGAAAAGCTGGGCAATTTGAATGCTTAATAGG	4134
1865	GCTTGTCTTCAGTGCAGTCTTACAAGGACACTTTAAAAAGAGATTGTCCAAGTAGAATAAG	1924
4135	GCTTGTCTTCAGTGCAGTCTTACAAGGACACTTTAAAAAGAGATTGTCCAAGTAGAATAAG	4194
1925	CCGCCCTTTCGTTCATGCCCTTTATTTCAGGGAATCACTGGAGGCCCACTGCCCCAGG	1984
4195	CCGCCCTTTCGTTCATGCCCTTTATTTCAGGGAATCACTGGAGGCCCACTGCCCCAGG	4254
1985	GGACAAAGTCTTTTGTAGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	2044
4255	GGACAAAGTCTTCTGTAGTCAGAAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	4314
2045	TGCTTGGGGAAGCGCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTTGACC	2104
4315	TGCTTGGGGAAGCGCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTTGACC	4374
2105	ATTGAGGGCCAGGAAGTTGCTCTCTGGACACTGGTGGCGGTCTTCTTACTTACTCTTTC	2164
4375	ATTGAGGGCCAGG-AGGTGCTCTCTGGACACTGGTGGCGGTCTTCTTACTTACTCTTTC	4433

Qy	2165	TGTC	CCGGA	CAACTGTCTCT	CCAGATCTGTCA	TATTTCTGAGGGGGTCC	NTAAGACGGGCA	2224
Db	4434	TGTC	CCGGA	CAACTGTCTCT	CCAGATCTGTCACTA	TCTGAGGGGGTCC	TAAGACGGGCA	4491
Qy	2225	GTCA	CTAGATACTTT	YTC	CCAGCCACTAAGTTATGA	ACTGGGAGCTTTATCTTTTCAC	2284	
Db	4492	GTCA	CTAGATAC	TTCT	CCAGCCACTAAGTTATG	ACTGGGAGCTTTATCTTTTCAC	4549	
Qy	2285	ATGCT	TTCTTAATATGCT	TTGAAAGCCCACTACCT	TTGTTAGGAGGAGACATTTCTAGCAA	2344		
Db	4550	ATGCT	TTCTTAATATGCT	TTGAAAGCCCACTACCT	TTGTTAGGAGGAGACATTTCTAGCAA	4609		
Qy	2345	AAGC	AGGGGCCATTTATAC	CTGAACATAGGAGAGGAA	CAACCCGTTGTGTGTC	2404		
Db	4610	AAGC	AGGGGCCATTTATAC	CTGAACATAGGAGAGGAA	CAACCCGTTGTGTGTC	4668		
Qy	2405	CTTG	AGGAAAGGAATTAAT	CTCTGAAGTCTGGGCAACAG	AAGGACAAATATGACAGCCAAA	2464		
Db	4669	CTTG	AGGAAAGGAATTAAT	CTCTGAAGTCTGGGCAACAG	AAGGACAAATATGACAGCCAAA	4727		
Qy	2465	GAATG	CCCGCTCTGTTCTCAAGT	TAAACTAAAGGATTCAC	TCTCTCTTCCCTACCAAGGCA	2524		
Db	4728	GAATG	CCCGCTCTGTTCTCAAGT	TAAACTAAAGGATTCAC	TCTCTCTTCCCTACCAAGGCA	4787		
Qy	2525	GTAC	CCCTCTCAGACCCCAAGG	CCCCAACAGGATTTCCAAAAGAT	TGTTAAAGACTTTAAAGC	2584		
Db	4788	GTAC	CCCTCTCAGACCCCAAGG	CCCCAACAGGATTTCCAAAAGAT	TGTTAAAGACTTTAAAGC	4847		
Qy	2585	CCAA	GGCTTAGTAAAAACCATG	CAATACTCCCTGCAGTAAT	TCGTTAGTGAGATGAGGAGG	2644		
Db	4848	CCAA	GGCTTAGTAAAAACCATG	CAATACTCCCTGCAGTAAT	TCGTTAGTGAGATGAGGAGG	4900		
Qy	2645	CACAGA	AAACCCAGTGACAGT	GGAGGTAGTCCAAAGAT	TCTCAGGATTAATCAATGAGGC	2704		
Db	4901	TACAGA	AAACCCAGTGACAGT	GGAGGTAGTCCAAAGAT	TCTCAGGATTAATCAATGAGGC	4958		
Qy	2705	CGTTG	TCTTTTATACCCAGCTG	ACTAGCCCTTATAC	TGTGMYTTC	CCAAATACAGA	2764	
Db	4959	TGTTG	TCTCTATAGCCAGCTG	ACTAGCCCTTATAC	TCTGCTTTCC	CAATACAGA	5018	
Qy	2765	GGAAG	CAGAGTGGTTTACAST	CTCTGCACTT	WAGGATGCTTCTTCTG	CAATCCCTGTACA	2824	
Db	5019	GGAAG	CAGAGTGGTTTACAGT	CTCTGCACTT	CAGATGCTTCTTCTG	CAATCCCTGTACA	5078	
Qy	2825	TCCTG	ACTCTCAATCTTGTT	GGCTTCAAGATCTTCA	AAACCCACATCTCAACTCAC	2884		
Db	5079	TCCTG	ACTCTCAATCTTGTT	GGCTTCAAGATCTTCA	AAACCCACATCTCAACTCAC	5138		
Qy	2885	CTGG	ACTTTTTACCCCAAGGGT	TCAGGGATAGYCCCCAT	CTATTTGGCCAGGCATTAGC	2944		
Db	5139	CTGG	ACTTTTTACCCCAAGGGT	TCAGGGATAGTCCCCAT	CTATTTGGCCAGGCATTAGC	5198		
Qy	2945	CCAAG	ACTTGAGYCAATYMTCA	TACTGCACTCTTGTC	TCCTCRGTAKGTGGATGATTT	3004		
Db	5199	CCAAG	ACTTGAGCAATCTCTCA	TACTGCACTCTTGTC	TCCTCRGTAKGTGGATGATTT	5256		
Qy	3005	ACTTT	TGCGCCYRTTCAGAA	ACCTTGGCCATC	AGCCACCCAGC	CBCTCTTAAATTT	3064	
Db	5257	ACTTT	TGCGCGCCATTCAGAA	ACCTTGGCCATC	AGCCACCCAGC	CBCTCTTCAATTT	5316	
Qy	3065	CTCG	CYACTGTGGCTAC	AWGGTTTCCAAAG	CSARAGCTCA	CTCTGCTCACAGCAGT	3124	
Db	5317	CTCG	CYACTGTGGCTACAT	GGTTTCCAAACCA	AAAGGCTCA	CTCTGCTCACAGCAGT	5376	
Qy	3125	TAAAT	ACTTAGGCTTAARA	TTATCCAAAGGCA	CCAGGGCCCTCAGT	GAGGAAYRATCCA	3184	
Db	5377	----	TACTTAGGCTAAAAAT	TATCCAAAGGCA	CCAGGGCCCTCAGT	GAGGAACACATCCA	5432	
Qy	3185	GCCT	TACTGGCTTATTCCT	CATCYCAAAAC	CCCTTAAAGCA	CTAAGRRHTTCTCTGGGCT	3244	
Db	5433	GCCT	TACTGGCTTATTCCT	CATCCAAA	CCCTTAAAGCA	CTAAGGGATCTCTGGGCT	5492	





Db	7636	TTAAAGTGGCTGGAGTCTTGGGATCATCACACTTGAGTCAAACTCTGGATCTGC	7695
Qy	5464	CAAGAGAACCTGAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGATT	5523
Db	7696	CNAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGATT	7755
Qy	5524	TGGCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACATAAATCAATAATCCCCCATG	5583
Db	7756	TGGCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACATAAATCAATAATCCCCCATG	7814
Qy	5584	GSCTCCCTTATCATATTTTCTCTKTASTGTSTTYYACCTSTTTTCACTCTCATGCA	5643
Db	7815	GCCTCCCTTATCATATTTTCTCTKTASTGTSTTYYACCTSTTTTCACTCTCATGCA	7874
Qy	5644	CCCTCTCATGCGCTGTATGACCAGTAGTCCCTCYACWAGAGTTTCTATGAGGAATG	5703
Db	7875	CCCTCTCATGCGCTGTATGACCAGTAGTCCCTCYACWAGAGTTTCTATGAGGAATG	7934
Qy	5704	CAGCGTCCCGAAATATTGATGCCCAATCGTATAGGAGTCTTTTSTAAGGGAACCCCAACC	5763
Db	7935	CAGCGTCCCGAAATATTGATGCCCAATCGTATAGGAGTCTTTTSTAAGGGAACCCCAACC	7994
Qy	5764	TTCACTGCCACACCCCATATGCCCCGCAACTGTGTATCACTCTGCGCACTCTTTGATGCAT	5823
Db	7995	TTCACTGCCACACCCCATATGCCCCGCAACTGTGTATCACTCTGCGCACTCTTTGATGCAT	8054
Qy	5824	GCAAACTACTATTATGGACAGGAAATAATGATTAATCCTAGTTGTCTCTGGAGACTTGA	5883
Db	8055	GCAAACTACTATTATGGACAGGAAATAATGATTAATCCTAGTTGTCTCTGGAGACTTGA	8114
Qy	5884	GTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTCTGATGGGGTGGAGTTCAA	5943
Db	8115	GTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTCTGATGGGGTGGAGTTCAA	8174
Qy	5944	GATCAGCAAGAGAAACAATGTAAGAAAGTAATCTCCAACTCACCCGGGTCAATGGC	6003
Db	8175	GATCAGCAAGAGAAACAATGTAAGAAAGTAATCTCCAACTCACCCGGGTCAATGGC	8234
Qy	6004	ACCTCTAGCCCTTACAAAGACTAGATCTCTGAAACTACATGAAACCCCTCCGTACCCAT	6063
Db	8235	ACCTCTAGCCCTTACAAAGACTAGATCTCTGAAACTACATGAAACCCCTCCGTACCCAT	8294
Qy	6064	ACTCGCTGTGAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAA	6123
Db	8295	ACTCGCTGTGAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAA	8354
Qy	6124	AACCTCTAACTGTGTGGATATGCTTCCCTGAACTTCARGCCATATGTTCAATCCCT	6183
Db	8355	AACCTCTAACTGTGTGGATATGCTTCCCTGAACTTCARGCCATATGTTCAATCCCT	8414
Qy	6184	GTACCTGAACAATGGGAACACTTCAGCACAGAAATAAACACCACTTCCGTTTATGAGGA	6243
Db	8415	GTACCTGAACAATGGGAACACTTCAGCACAGAAATAAACACCACTTCCGTTTATGAGGA	8474
Qy	6244	CCTCTGTGTTTCAATSTGGAATAAACCCATACCTCAAACTCACTGTGTGTAATTTAGC	6303
Db	8475	CCTCTGTGTTTCAATSTGGAATAAACCCATACCTCAAACTCACTGTGTGTAATTTAGC	8534
Qy	6304	AATACATACACAAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCCAACAATA	6363
Db	8535	AATACATACACAAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCCAACAATA	8594
Qy	6364	GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAAT	6423
Db	8595	GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAAAT	8654
Qy	6424	GGCTCTTCAAGATCTATGTCTCTCATCTTGTAGTCCGCCATGRCATCTACACT	6483
Db	8655	GGCTCTTCAAGATCTATGTCTCTCATCTTGTAGTCCGCCATGRCATCTACACT	8714
Qy	6484	GAAACAAGATTATACAGTTATGTATATCTTAAGCCCGCAACAAAAGAGTACCATTCTT	6543
Db	8715	GAAACAAGATTATACAGTTATGTATATCTTAAGCCCGCAACAAAAGAGTACCATTCTT	8774
Qy	6544	CTTTTGTATTATAGGACAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCACA	6603
Db	8775	CTTTTGTATTATAGGACAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCACA	8834
Qy	6604	ACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC	6663
Db	8835	ACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC	8894
Qy	6664	GCGACTCCCTGGTCACTTGTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCTTCA	6723
Db	8895	GCGACTCCCTGGTCACTTGTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCTTCA	8954
Qy	6724	AATCGAAGAGCTTTAGACTTGTCTTAACTCCGCTGAAAGAGGGGAAACCTGTTTATTTAGGG	6783
Db	8955	AATCGAAGAGCTTTAGACTTGTCTTAACTCCGCTGAAAGAGGGGAAACCTGTTTATTTAGGG	9014
Qy	6784	GAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	6843
Db	9015	GAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	9074
Qy	6844	CSAGATCGAATACAAAGTAKAGCAGARGAGCTTTCGAAAACACTGGAACCTTGGGGCTCTC	6903
Db	9075	CSAGATCGAATACAAAGTAKAGCAGARGAGCTTTCGAAAACACTGGAACCTTGGGGCTCTC	9134
Qy	6904	AGCCATGATGCTGCTGGAATCTCCCTCTTCTTAGGACCTCTAGCAGCTATATTTGCTA	6963
Db	9135	AGCCATGATGCTGCTGGAATCTCCCTCTTCTTAGGACCTCTAGCAGCTATATTTGCTA	9194
Qy	6964	CTCTCTTTGGACCTCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTCCAGAACTGAA	7023
Db	9195	CTCTCTTTGGACCTCTGTATCTTTTACCTCTCTTGTAACTTTGTCTCTCCAGAACTGAA	9254
Qy	7024	GCTGTAAACTCAAAATGGAGCCCAAGATGCAAGTCCAACTAAGATCTACCCAGACCC	7083
Db	9255	GCTGTAAACTCAAAATGGAGCCCAAGATGCAAGTCCAACTAAGATCTACCCAGACCC	9314
Qy	7084	CTGGAACCGCTGTAGCCCAAGTGTATGATGATGATGATGATGATGATGATGATGATGAT	7143
Db	9315	CTGGAACCGCTGTAGCCCAAGTGTATGATGATGATGATGATGATGATGATGATGATGAT	9374
Qy	7144	GAATCTCAGCTGCAACACTCTACTACGCCCAATTTACGAGGAAGAGTGTAGAGCGGT	7203
Db	9375	GAATCTCAGCTGCAACACTCTACTACGCCCAATTTACGAGGAAGAGTGTAGAGCGGT	9434
Qy	7204	SGTCGGCCAACTCTCCCAACAGCACTTAGGTTTCTGTGATGATGATGATGATGATGATGAT	7263
Db	9435	SGTCGGCCAACTCTCCCAACAGCACTTAGGTTTCTGTGATGATGATGATGATGATGATGAT	9494
Qy	7264	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTAGTGGGAGGTGACC	7323
Db	9495	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTAGTGGGAGGTGACC	9554
Qy	7324	ACATCCACTTTAAACAACCGGGCTTGCACCTTAGTCTCAACCTGACCAATCAGAGAGCTC	7383
Db	9555	ACATCCACTTTAAACAACCGGGCTTGCACCTTAGTCTCAACCTGACCAATCAGAGAGCTC	9614
Qy	7384	ACTAAATGCTAATTTAGGCAAGAGAGGTTAAAGAAATAGCCAAATCATYTTATTCGWTG	7443
Db	9615	ACTAAATGCTAATTTAGGCAAGAGAGGTTAAAGAAATAGCCAAATCATYTTATTCGWTG	9674
Qy	7444	AGAGCACAGCAGGAGGCAATGATCGGGATATAACCAAGTTCGAGCGGCAACGG	7503
Db	9675	AGAGCACAGCAGGAGGCAATGATCGGGATATAACCAAGTTCGAGCGGCAACGG	9734
Qy	7504	CAACCCCTTTGGGTCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT	7563
Db	9735	CAACCCCTTTGGGTCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT	9794
Qy	7564	ATTAAATCTTGCACTGCR	7582
Db	9795	ATTAAATCTTGCACTGCA	9813

RESULT 7	AX329572	AX329572	56093 bp	DNA	linear	PAT 09-JAN-2002
LOCUS	AX329572	Sequence 81 from Patent WO0194629.				
DEFINITION	AX329572	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
ACCESSION	AX329572.1	GI:18102550				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
ORIGIN						
Query Match		85.0%;	Score 6448;	DB 6;	Length 56093;	
Best Local Similarity		95.7%;	Pred. No. 0;			
Matches 6681;		Conservative 195;	Mismatches 50;	Indels 53;	Gaps 27;	
QY	606	GATGGGAAACGTTCCCGCAAGACAAACGCGCCCTAAGACGCTATTCTGGGAAATTTGGGA	665			
DB	30953	GATGGGAAACGTTCCCGCAAGACAAACGCGCCCTAAGACGCTATTCTGGGAAATTTGGGA	31012			
QY	666	MCAATTGTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTGGCGCTG	725			
DB	31013	CCAATTGTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCGAGTGGCGCTG	31072			
QY	726	GCACCTCTGAGGAAGTATATATTAACACCATCTTACAGCTAGACACTCTTTTGTAGAA	785			
DB	31073	GCACCTCTGAGGAAGTATATATTAACACCATCTTACAGCTAGACACTCTTTTGTAGAA	31132			
QY	786	AAGGCAATGGAGTGAAGTGCCATAAGTACAAACTTCTTTTCAATAAGACAACTCAC	845			
DB	31133	AAGGCAATGGAGTGAAGTGCCATAAGTACAAACTTCTTTTCAATAAGACAACTCAC	31192			
QY	846	AATTATGTAAAAAGTGTGATTATGCGCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	905			
DB	31193	AATTATGTAAAAAGTGTGATTATGCGCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	31252			
QY	906	CGAGCATCCCGGACTCTTCCCAATTAATAGGACCCCTTCAACCAATGGTCCAA	965			
DB	31253	CGAGCATCCCGGACTCTTCCCAATTAATAGGACCCCTTCAACCAATGGTCCAA	31312			
QY	966	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATATGA	1025			
DB	31313	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATATGA	31372			
QY	1026	CCCTCCCAAGCAGTGGAGGAAGAAATTCGGGCCCGCAGAGTGCATGTCGCTTTTYY	1085			
DB	31373	CCCTC-CCAAGCAGTGGAGGAAGAAATTCGGGCCCGCAGAGTGCATGTCGCTTTTTC	31431			
QY	1086	TCTCCGAGCTTAAGCAATTAACACAGACTTAGTAAATTTCTCAGATAAATCTGATGG	1145			
DB	31432	TCTCCGAGCTTAAGCAATTAACACAGACTTAGTAAATTTCTCAGATAAATCTGATGG	31491			
QY	1146	CTATATTGRTGTTTACAAAGGGTTAGGACAAATTTCTTTGATCTGCAATGGAGATATATA	1205			
DB	31492	CTATATTGRTGTTTACAAAGGGTTAGGACAAATTTCTTTGATCTGCAATGGAGATATA	31550			
QY	1206	TGTCACCTGCTAAATTCAGACACTAACCCCAATAGAGAGAAAGTGCACCAATCTGAGCCT	1265			
DB	31551	TGTCACCTGCTAAATTCAGACACTAACCCCAATAGAGAGAAAGTGCACCAATCTGAGCCT	31610			



Db 34803 CCAATGGAAATTACTTAAACCCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACC 34862  
Qy 4564 TCARATGCCAAATCATTATTACTTGGACCAAGGCTTTTCAAAAATCAATCAAGCARATAKT 4623  
Db 34863 TCAGATGCCAAATCATTATTACTTGGACCAAGGCTTTTCAAAAATCAATCAAGCARATAGT 34922  
Qy 4624 CAGGCGCTGTGAATGTGTCARABAAAATAATCCCTGCTGCTTATCGCCAAAGCTCTTCAGG 4683  
Db 34923 CAGGCGCTGTGAATGTGTCARABAAAATAATCCCTGCTGCTTATCGCCAAAGCTCTTCAGG 34982  
Qy 4684 ARAACAARAACAGGCCATTCACCTGRRARARACTGGCACTGATTTTACCACAAGGCC 4743  
Db 34983 AGAACAAAGAACAGGCCATTCACCTGGAGAAAGACTGGCACTGATTTTACCACAAGGCC 35042  
Qy 4744 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGGTGGGCARAG 4803  
Db 35043 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGGTGGGCARAG 35102  
Qy 4804 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAATA 4863  
Db 35103 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAATA 35162  
Qy 4864 ATTCCGAGATTCGGAATTCCTCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGGCC 4923  
Db 35163 ATTCCGAGATTCGGAATTCCTCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGGCC 35222  
Qy 4924 ACAGTAACCCAGGAGTATCCAGGCTTGTAGTATACGATATCACTTACACTGGCCCTGA 4983  
Db 35223 ACAGTAACCCAGGAGTATCCAGGCTTGTAGTATACGATATCACTTACACTGGCCCTGA 35282  
Qy 4984 AGGCCACAGTCTCAGGGAAGGTCCAGAAAATGAAATGAAAYACTCAAAAGGACATCTAAA 5043  
Db 35283 AGGCCACAGTCTCAGGGAAGGTCCAGAAAATGAAATGAAACACTCAAAAGGACATCTAAA 35342  
Qy 5044 AAGCAAAACCCAGGAAACCCACTCATGCGCTGTCTGTGCTATAGCCTTAAAGA 5103  
Db 35343 AAGCAAAACCCAGGAAACCCACTCATGCGCTGTCTGTGCTATAGCCTTAAAGA 35402  
Qy 5104 ATCTGCAATCTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 5163  
Db 35403 ATCTGCAATCTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 35462  
Qy 5164 TTCTAAACCAATGACTGTGCTGCACCCAGACAGCCAACTTAGTGTGACAGATCACT 5223  
Db 35463 TTCTAAACCAATGACTGTGCTGCACCCAGACAGCCAACTTAGTGTGACAGATCACT 35522  
Qy 5224 CTTTAGCCAAATATCAACAAGTTCTTAAACCAATTAAGGAACCTTATCCCTGAGAAGAG 5283  
Db 35523 CTTTAGCCAAATATCAACAAGTTCTTAAACCAATTAAGGAACCTTATCCCTGAGAAGAG 35582  
Qy 5284 GAAGAAGACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC 5343  
Db 35583 GAAGAAGACTATTCACCCCTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAATTC 35642  
Qy 5344 CATCCCTAGATACATCTGGGAAGGACCTACCCAGTCAATTTATYTATACCCCACTGCG 5403  
Db 35643 CATCCCTAGATACATCTGGGAAGGACCTACCCAGTCAATTTATCTACCCCACTGCG 35702  
Qy 5404 TTAAGTGGCTGGATGGAGTCTTGATACATCAACTTGAAGTCAAAATCTCGGATCTGC 5463  
Db 35703 TTAAGTGGCTGGATGGAGTCTTGATACATCAACTTGAAGTCAAAATCTCGGATCTGC 35762  
Qy 5464 CAAAGGAACCTGAAATCCAGAGACAGCTTAGTATTCCTGTAACCTCTAGAGGATT 5523  
Db 35763 CAAAGGAACCTGAAATCCAGAGACAGCTTAGTATTCCTGTAACCTCTAGAGGATT 35822  
Qy 5524 TCGCGCTGCTTCAAAACAAACAGAGGAAAGTAACTAAATCATAAATCCCCCATG 5583  
Db 35823 TCGCGCTGCTTCAAAACAAACAGAGGAAAGTAACTAAATCATAAATCCCCCATG 35881  
Qy 5584 GSCCTCCCTATCATATTTTCTCTKNTASTGTSTTTTACCTSTTTCACTCTCACTGCA 5643  
Db 35882 GSCCTCCCTATCATATTTTCTCTTACTGTTCTTTTACTGTTCTTTTACTCTCTCACTGCA 35941

Qy 5644 CCCCCTCCATGCGCTGTATGACCCAGTAGTCCCTTACCMAGAGTTTCTTATGAGAAATG 5703  
Db 35942 CCCCCTCCATGCGCTGTATGACCCAGTAGTCCCTTACCAAGAGTTTCTTATGAGAAATG 36001  
Qy 5704 CAGGCTCCGGAATAATTTGATGCGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAACC 5763  
Db 36002 CAGGCTCCGGAATAATTTGATGCGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAACC 36061  
Qy 5764 TTCACTGCCCAACCCCATATGCCCCGCAATGCTCTATCACTCTGCACTCTTTTGATGCAAT 5823  
Db 36062 TTCACTGCCCAACCCCATATGCCCCGCAATGCTCTATCACTCTGCACTCTTTTGATGCAAT 36121  
Qy 5824 GCAATACTCATTTATGCGACAGGAAAAATGATTAATCCTAGTTGCTCTGGAAGGACTTGA 5883  
Db 36122 GCAATACTCATTTATGCGACAGGAAAAATGATTAATCCTAGTTGCTCTGGAAGGACTTGA 36181  
Qy 5884 GTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTCAA 5943  
Db 36182 GTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTCAA 36241  
Qy 5944 GATCAGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCAACTCACCCGGGTACATGGC 6003  
Db 36242 GATCAGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCAACTCACCCGGGTACATGGC 36301  
Qy 6004 ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAAATACATGAAGAACCTCCGTACCCAT 6063  
Db 36302 ACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAAATACATGAAGAACCTCCGTACCCAT 36361  
Qy 6064 ACTCGCTGTGAAGCCTATTTAATACCACTCTGATGGGCTCCATGAGGTCTCGGCCAA 6123  
Db 36362 ACTCGCTGTGAAGCCTATTTAATACCACTCTGATGGGCTCCATGAGGTCTCGGCCAA 36421  
Qy 6124 AACCCTACTACTGTGATATGCTCCCTGTAATCTTCAAGCCATATGTTTCAATCCCT 6183  
Db 36422 AACCCTACTACTGTGATATGCTCCCTGTAATCTTCAAGCCATATGTTTCAATCCCT 36481  
Qy 6184 GTACTGAACAAATGGAAACACTTCAGCAAGAAATAAACACCACTTCCTGTTTATGAGGA 6243  
Db 36482 GTACTGAACAAATGGAAACACTTCAGCAAGAAATAAACACCACTTCCTGTTTATGAGGA 36541  
Qy 6244 CCTCTGTTTCCAAATSTGGAAATAACCATACCTTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 36542 CCTCTGTTTCCAAATSTGGAAATAACCATACCTTCAAACTCACCTGTGTAAATTTAGC 36601  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCAATCAGTGGGTAACTCTCCCAACAAATA 6363  
Db 36602 AATACTACATACAAACCAACTCCCAATGCAATCAGTGGGTAACTCTCCCAACAAATA 36661  
Qy 6364 GTCTGCTTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCGTGTGTTGAAT 6423  
Db 36662 GTCTGCTTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCGTGTGTTGAAT 36721  
Qy 6424 GGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTCCATGRCATCTACACT 6483  
Db 36722 GGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTTATGACCATCTACACT 36781  
Qy 6484 GAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTAGTACCTTCTT 6543  
Db 36782 GAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTAGTACCTTCTT 36841  
Qy 6544 CCTTTTGTATAGGACAGGAGTGTAGTGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 6603  
Db 36842 CCTTTTGTATAGGACAGGAGTGTAGTGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 36901  
Qy 6604 ACCTCTACTCAGTTCTACTAATACTATCTAAGAACTAAATGSSGACATGGAACGGGTC 6663  
Db 36902 ACCTCTACTCAGTTCTACTAATACTATCTAAGAACTAAATGSSGACATGGAACGGGTC 36961  
Qy 6664 GCGGACTCCCTGGTCCAGACTCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCA 6723  
Db 36962 GCGGACTCCCTGGTCCAGACTCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCA 37021







QY	966	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATCCCAATATATGA	1025	Db	32382	TGCTGGGGCAAGGCCATCCATGCGATCAACCTCCACAGAGCCCTGGGTATGCTTGACC	32441
Db	31313	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATCCCAATATATGA	31372	QY	2105	ATTGAGGGCCAGGAAGGTGTCTCTCGACACTGCTGGGCTCTTTAGTCTTACTCTTC	2164
QY	1026	CCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGGAGTGCATGTCTTTT	1085	Db	32442	ATTGAGGGCCAGG-AGGTTGTCTCTCGACACTGCTGGGCTCTTCTTAGTCTTACTCTTC	32500
Db	31373	CCCT-CCAAAGCAGTGGGAGGAGAAATTCGGCCCGCAGGAGTGCATGTCTTTT	31431	QY	2165	TGTCGCCGACAACTGTCTCCAGATCTCTCACTATCTTGAGGGGTCCNTAAGACGGCA	2224
QY	1086	TCTCCAGACTTAAAGCAATAAACAAGACTTAAAGTAAATCTCAGATAAATCCCTGATGG	1145	Db	32501	TGTCGCCGACAACTGTCTCTCAGATCTCTCACTA-TCTGAGGGGTCC-TAAGACGGCA	32558
Db	31432	TCTCCAGACTTAAAGCAATAAACAAGACTTAAAGTAAATCTCAGATAAATCCCTGATGG	31491	QY	2225	GTCACTAGATCTTTTCTCCAGCCACTAAGTTATGAATCTGGGAGCTTATCTTTTAC	2284
QY	1146	CTATATGRTGTTTAAAGGTTAGGACAAATCTTTTGTATCTGACATGAGAGATATATA	1205	Db	32559	GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGAGCTTATCTTTTAC	32616
Db	31492	CTATATGATGTTTAAAGGTTAGGACAAATCTTTTGTATCTGACATGAGAGATATA-A	31550	QY	2285	ATGCTTTTCTAATATGCTTGAAGCCCACTACCTGTTAGGAGAGACATCTTAGCAA	2344
QY	1206	TGTCACCTGTAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAATCTGAGCCT	1265	Db	32617	ATGCTTTTCTAATATGCTTGAAGCCCACTACCTGTTAGGAGAGACATCTTAGCAA	32676
Db	31551	TGTCACCTGTAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAATCTGAGCCT	31610	QY	2345	AAGCAGGGGCCATATACACCTGACATAGGAGAGAACACCCGTTTGTGTCCTG	2404
QY	1266	GAGGTGTTGGGATCTCTGTTATCTCAGTCAAGTCAATGATPANGATGACAAACAGAGG	1325	Db	32677	AAGCAGGGGCCATATACACCTGACATAGGAGAGAACACCCGTTTGTGTCCTG	32735
Db	31611	GAGGTGTTGGGATCTCTGTTATCTCAGTCAAGTCAATGATPANGATGACAAACAGAGG	31667	QY	2405	CTTGAGGAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGGACATATGGACGACCAA	2464
QY	1326	AAAGANAATGATTCCTCCACAGCCAGCAGCAGTTCCTCAGTCTAGACCTCAATCTGGGAC	1385	Db	32736	CTTGAGGAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGGACATATGGACGAG-CAAA	32794
Db	31668	AAAGANAATGATTCCTCCACAGCCAGCAGCAGTTCCTCAGTCTAGACCTCAATCTGGG	31724	QY	2465	GAATGCCGCTCTGTTCAAGTTAACTAAAGGATTCCTCTTCCCTACCAAGGCA	2524
QY	1386	ACAGNAATCAGTAACATGGGAGATTTGGTCTGCAGACATTTGCTAACTTTGTGTASAA	1445	Db	32795	GAATGCCGCTCTGTTCAAGTTAACTAAAGGATTCCTCTTCCCTACCAAGGCA	32854
Db	31725	ACACAGAAATCAGAACT-GGAGATTTGGTCTGCAGACATTTGCTAACTTTGTGTAGAA	31783	QY	2525	GTACCCCTCAGAGCCAGGCCCACAGGATTCCTAAAGATTTGTTAGGACTTAAAGC	2584
QY	1446	GGACTAAGGAAACTTASGAAGAAATCTAYGAATTTACTCAATGATGTCACCATACACA	1505	Db	32855	GTACCCCTCAGAGCCAGGCCCACAGGATTCCTAAAGATTTGTTAGGACTTAAAGC	32914
Db	31784	GGACTAAGGAAACTTASGAAGAAATCTAYGAATTTACTCAATGATGTCACCATACACA	31842	QY	2585	CCAAGGCTTAGTAAACCATGCAATACTCCCTGCAAGTAACTCCGTAGTGGATTGAGGAGG	2644
QY	1506	GGGAAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTTAGGAGAGC	1565	Db	32915	CCAAGGCTTAGTAAACCATGCAATACTCCCTGCAAGTAACTCCGTAGTGGATTGAGGAGG	32967
Db	31843	GGGAAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTTAGGAGAGC	31901	QY	2645	CACAGAAACCCAGTGGAGCAGTGGAGGTTAGTGAAGATTCAGAGATTCATGAGAGC	2704
QY	1566	GTGCTCTCTGCTCAGCTCTTCTGAAGGCAACTAATCTTAAAGGTTAAGTTATCA	1625	Db	32968	TACAGAAACCCAAACAGACAGTGA-GGTTAGTGAAGATCTCAGGATTCAT-AGGC	33025
Db	31902	GTGCTCTCTGCTCAGCTCTTCTGAAGGCAACTAATCTTAAAGGTTAAGTTATCA	31961	QY	2705	GGTGTCTCTTTATACCCAGCTGACCTAGCCCTTATCTGTMVTTCCCAATACCCAGA	2764
QY	1626	CTCAGTCACTGTCAGACATTAG-AAAGAACTTCAAAAGTCTGCGTAGGCGCCGAGCAAA	1684	Db	33026	TGTTGTCTCTATAGCAGCTGACCTAGCCCTTATCTGCTCTTCCCAATACCCAGA	33085
Db	31962	CTCAGTCACTGTCAGACATTAGAAAGAACTTCAAAAGTCTGCGTAGGCGCCGAGCAAA	32021	QY	2765	GGAGCAGAGTGGTTACASTCCTGGACCTTMAAGATGCTTCTGTCATCCCTGTATCA	2824
QY	1685	ACTTAGAAACCTATTGAACTTGGCAACTGCGTTTATTAATAGAGATCAGGAGGAGC	1744	Db	33086	GGAGCAGAGTGGTTACAGTCTGACCTTCAAGATGCTTCTGATCCCTGTATCA	33145
Db	32022	ACTTAGAAACCTATTGAACTTGGCAACTGCGTTTATTAATAGAGATCAGGAGGAGC	32081	QY	2825	TCTGACTCTCAATCTTCTGCTTTGAGATTAAGATTAAGATTAAGATTAAGATTAAG	2884
QY	1745	AGGCGGAACAGCAACCGGATTAAGAAAGGCAACCGCTTTAGTCAATGACCTCAGG	1804	Db	33146	TCTGACTCTCAATCTTCTGCTTTGAGATTAAGATTAAGATTAAGATTAAGATTAAG	33205
Db	32082	AGGCGGAACAGCAACCGGATTAAGAAAGGCAACCGCTTTAGTCAATGACCTCAGG	32141	QY	2885	CTGACTCTTTTACCCCAAGGGTTCAGGATAGTCCCATCTATTTGGCCAGGCTTAGC	2944
QY	1805	CAAGTGACTTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAATTTGAATCCCTAATAGG	1864	Db	33206	CTGACTCTTTTACCCCAAGGGTTCAGGATAGTCCCATCTATTTGGCCAGGCTTAGC	33265
Db	32142	CAAGTGACTTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAATTTGAATCCCTAATAGG	32201	QY	2945	CCAAGACTTGAGYCATYMTCACTCTGGACACTCTTGTCTCTCRGTAKGTGGATGATTT	3004
QY	1865	GCTTGTCTTCAAGTGGGCTTACAAGGACACTTTAAAGAAAGATTTGCAAGTAGAAGTAAG	1924	Db	33266	CCAAGACTTGAGYCATYMTCACTCTGGACACTCTTGTCTCTCRGTAKGTGGATGATTT	3064
Db	32202	GCTTGTCTTCAAGTGGGCTTACAAGGACACTTTAAAGAAAGATTTGCAAGTAGAAGTAAG	32261	QY	3005	ACTTTTGGCGCTTTTCAAGAAACCTTTGTCATCAAGCCCAACCAAGCTCTTTTAAATTT	3064
QY	1925	CCGCCCCCTTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAAGGCCACTGCCCCAGG	1984	Db	33324	ACTTTTGGCGCTTTTCAAGAAACCTTTGTCATCAAGCCCAACCAAGCTCTTTTAAATTT	33383
Db	32262	CCGCCCCCTTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAAGGCCACTGCCCCAGG	32321	QY	3065	CCTCGCTGCTGCTTACAGGCTTTTCCAAACCSARGCTCARTCTCTGCTCAGAGCAGGT	3124
QY	1985	GGACAAAGGCTTTTGTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGCTGAGG	2044	Db	33384	CCTCGCTGCTGCTTACAGGCTTTTCCAAACCSARGCTCARTCTCTGCTCAGAGCAGGT	33443
Db	32222	GGACAAAGGCTTTTGTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGCTGAGG	32381	QY	3125	TAAATCTTAGGCTTAARATTTATCCAAAGGCAACCCAGGCTCCTGAGTGGAGAAATATCA	3184
QY	2045	TGCTGGGGCAAGGCCCATCCATGCTCCATCCTCAGAGCCTCTGGGTATGCTTGACC	2104				

Db	33444	-----TACTTTAGGCGTAAAAATTATCCAAAGGCCACGAGGCCCTCAGTGGAGAACACATCCA	33499	QY	4265	GGAAAGRAAAGGGAAATASAAAGRAAYTGCCAAAGCAKATATTTGAAGCMAAAGAGCTGCA	4324
QY	3185	GCCTACTGGCTTATCTTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTTGGGRT	3244	Db	34563	GGAAAGGAAAGGGAAATAGAAGAACTGCCAAGCAGATATTGAAGCCAAAGAGCTGCA	34622
Db	33500	GCCTACTGGCTTATCTTCATCCCAAAACCCCTAAAGCAACTAAGGGGNTTCTTTGGGCT	33559	QY	4325	AGGCAAGGACCTCCATTAGAAATGCTTTATAAACTTCCCTTAGTATAGGTAATCCCTTC	4384
QY	3245	AAAYAGYTTCTCCGGAAMATGATTTCCCCAGGTGTGGCRRAAATAGCCAGGYCATTAWATA	3304	Db	34623	AGGCAAGGACCTCCATTAGAAATGCTTTATAAAACCAACCCCTAGTATAGGTAATCCCTTC	34682
Db	33560	ATATAGTTTCTGCGGAATGATTT-CCACAGTATGGCGAATATAGCCAGTCAATTAATA	33618	QY	4385	CGGGAACCAAGCCCCAGTACTCAGCAGAGAAACAGAAATGGGAACCTCAGCAGG-CAG	4443
QY	3305	CASTAATTAAGGAACACTCAGAAGCCAATACCAATTTARTATAGATGGAYAMCTGAAGYMR	3364	Db	34683	CGGGAACCAAGCCCCAGTACTCAGCAGAGAAACAGAAATGGGAACCTCAGCAGAGACAG	34742
Db	33619	CACATAATTAAGGAACACTCAGAAGCCAATACCAATTTAGTAAGATGGACAACATGAAGTAG	33678	QY	4444	TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGGAAAAATACATTTTTCCTGCAACTAT	4503
QY	3365	AGTGGCTTTCCAGGCCCCCTAAGAAGGCCCTTAACCCCAAGYCCCAAGTGTAAAGYTTGCC	3424	Db	34743	TTTTCTCCCTCGGACGCTTAGCCACTGAAGAAGGGAAAAATACATTTTTCCTGCAACTAT	34802
Db	33679	AAAGTGGCTTTCCAGGCC-----TAACCCCAAGCCCCAGTGTAAAGYTTGCC	33724	QY	4504	CCATAGGAATTAATTAATAACCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCACCA	4563
QY	3425	AACRGGGCAAGACTTTTSTTYATVRTACAGAAAAAAACAGRAAYAGCTCTRGAGTCC	3484	Db	34803	CCATAGGAATTAATTAATAACCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCACCA	34862
Db	33725	AACAGGGCAAGACTTTTTCATATGTACAG-AAAAACAGGAATAGCTCTAGGAGTCC	33783	QY	4564	TCARATGGCCAAATCATTTATTACTGACACGAGGCCCTTTTCAAACTATCAAGCARATAT	4623
QY	3485	TTACACAGTCCRAGGGAYGAGTTGCCAACCTRTGGCRVACCTGASTAAGGAAATYGATG	3544	Db	34863	TCAGATGGCCAAATCATTTATTACTGACACGAGGCCCTTTTCAAACTATCAAGCARATAT	34922
Db	33784	TTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCATACCTGACTAAGGAAATTTGATG	33843	QY	4624	CAGGGCCTGTGAATGTGCCARAAATAATCCCTGCTTATCCCAAGCTCTTTCAGG	4683
QY	3545	TAGTGGCAAGGTTGRCVYTCATTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT	3604	Db	34923	CAGGGCCTGTGAAGTGTGCCAGAGAAATAATCCCTGCTTATCCCAAGCTCTTTCAGG	34982
Db	33844	TAGTGGCAAGGTTGACCTCAATTTGTACGGGTAGTGTGGCAGTAGCAGTCTTAGTAT	33903	QY	4684	ARAAACAAAPACAGGCCATTTACCTGTRARAPACTGGCACTGATTTTACCCACAGGCC	4743
QY	3605	CTGAAGCAGTTAAATAATACAGGGGAGAGACTTACTGTGTGGAATCTCATGAKGTCA	3664	Db	34983	AGAAACAAAGAACAGGCCATTTACCTGGAGAGACTGGCACTGATTTTACCCACAGGCC	35042
Db	33904	CTGAAGCAGTTAAATAATACAGGGGAGAGACTTACTGTGTGGAATCTCATGATGTCA	33963	QY	4744	AAACCTCAGGGATTTTCACTGATCTACTAGTCTGGGTARATACCTTCAACGGGTTGGGCGARAG	4803
QY	3665	AYRGCACTACTGCTGTAAGGAGACTTGTGGCTGTGCAGAACAYGTTTACTTAATRTFC	3724	Db	35043	AAACCTCAGGGATTTTCACTGATCTACTAGTCTGGGTARATACCTTCAACGGGTTGGGCGAG	35102
Db	33964	ATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGCAGAACCTGTTTACTTAATGTC	34023	QY	4804	GCCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACACTAGTTTCAATAA	4863
QY	3725	AGGCTATATCTTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCACTTCTTAACCCAG	3784	Db	35103	GCCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACACTAGTTTCAATAA	35162
Db	34024	AGGCTATATCTTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCACTTCTTAACCCAG	34083	QY	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTTCCAGGCC	4923
QY	3785	YNCATTTCTTCAGACAATGAAGAAAGATARAAYATAACTGTCAACAAATAATTTCTC	3844	Db	35163	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTTCCAGGCC	35222
Db	34084	CCACATTTCTTCAGACATGAGAAAGATATAAACAATACTGTCAACAAAGTAATTTCTC	34143	QY	4924	ACAGTAACCCAGGAGTATCCCGAGGCTTACAGAGTATACATCTTACACTTACACTGCGCTGA	4983
QY	3845	AAACCTATGCCACTCGAGGGGACCTTGTAGARTTCCTTGTGACTGATCCYVAGCTTCAAC	3904	Db	35223	ACAGTAACCCAGGAGTATCCCGAGGCTTACAGAGTATACATCTTACACTTACACTGCGCTGA	35282
Db	34144	AAACCTATGCCACTCGAGGGGACCTTGTAGAGGTTCTTTAGAGGTTCTTTGACTGATCCCGACC-TCAAC	34202	QY	4984	AGGCCACAGTCTCTCAGGGAGGTTCGAGAAATGAATGAAYACTCAAGAGCATCTAATAA	5043
QY	3905	TTGTATAGTATGAGTCTCTTTGTAGAAAAAGGACTTCGAAAGYGGGGTATGAGTG	3964	Db	35283	AGGCCACAGTCTCTCAGGGAGGTTCGAGAAATGAATGAAYACTCAAGAGCATCTAATAA	35342
Db	34203	TTGTATAGTATGAGTCTCTTTGTAGAAAAAGGACTTCGAAAGYGGGGTATGAGTG	34262	QY	5044	AAGCAAAACCCAGGAACCCACCTCAGCATGGCTGTCTGCTGCTATAGCTTTAAAGA	5103
QY	3965	GTCAGTGATAATGGAATYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTAGCTR	4024	Db	35343	AAGCAAAACCCAGGAACCCACCTCAGCATGGCTGTCTGCTGCTATAGCTTTAAAGA	35402
Db	34263	GTCAGTGATAATGGAATYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTAGCTR	34322	QY	5104	ATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163
QY	4025	GCAGAACTAATAGCCVTCATYKGGGCACTAGAAATTAGGAGAAAGAAAGGGYAATATA	4084	Db	35403	ATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	35462
Db	34323	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAAAGAAAGGGCAAAATATA	34382	QY	5164	TTTCATAACCAATGACTTGTGCTTGACCCCAAGACAGCCAACTTAGTTTGCAGACATCACT	5223
QY	4085	TATACAGACTCTPARTATGCTYACCTAGTCTNTCCATGCCCATGMRGCCAATATGSARAGAA	4144	Db	35463	TTTCATAACCAATGACTTGTGCTTGACCCCAAGACAGCCAACTTAGTTTGCAGACATCACT	35522
Db	34383	TATACAGACTCTAATATATGCTTACCTTAGTCTCTCATGCCCATGAGCAATATGGAAGAA	34442	QY	5224	CCTTAGCCAAATATCAAGTTCTTTAAACATTACAAGAACCTATCCCTCGAGAGAGG	5283
QY	4145	AGGGAATTCCTAATCTCYAGRGAACACCTTATCAMAATCAGGAAGCCATTAGGARATTA	4204	Db	35523	CCTTAGCCAAATATCAAGTTCTTTAAACATTACAAGAACCTATCCCTCGAGAGAGG	35582
Db	34443	AGGGAATTCCTAATCTCYAGRGAACACCTTATCAMAATCAGGAAGCCATTAGGARATTA	34502	QY	5284	GAAGAAGCACTATTTCACCCWMTGACATGATATTAGTCAAGTCCCTTCYCTCTAATTTCCC	5343
QY	4205	TTATYGGCWGTACAGAAACCTTARAGAGGTGGMAGTCTTACCTGCGGGGTCTATCANAAA	4264	Db	35583	GAAGAAGCACTATTTCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTTCCC	35642
Db	34503	TTATYGGCWGTACAGAAACCTTARAGAGGTGGMAGTCTTACCTGCGGGGTCTATCANAAA	34562				



Db 37802 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCATGCTATTTCACCTCT 37861

Qy 7564 ATTAATCTTGCACCTGCR 7582

Db 37862 ATTAATCTTGCACCTGCA 37880

RESULT 9

AC007566/c

LOCUS AC007566 149194 bp DNA linear PRI 01-MAR-2002

DEFINITION Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.

ACCESSION AC007566

VERSION AC007566.2 GI:11181861

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 149194)

Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

99063792

9847074

2 (bases 1 to 149194)

Du, Z.

The sequence of Homo sapiens BAC clone CTB-10G5

Unpublished (2001)

3 (bases 1 to 149194)

Waterston, R.H.

Direct Submission

Submitted (15-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 149194)

Waterston, R.

Direct Submission

Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

5 (bases 1 to 149194)

Waterston, R.H.

Direct Submission

Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 149194)

Waterston, R.H.

Direct Submission

Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

7 (bases 1 to 149194)

Waterston, R.H.

Direct Submission

Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

8 (bases 1 to 149194)

Waterston, R.

Direct Submission

Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Nov 16, 2000 this sequence version replaced gi:4835815.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [sapiens@wustl.edu](mailto:sapiens@wustl.edu)

----- Summary Statistics

Center project name: H\_RG010G05

-----

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI McPherson 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelOBAcl1

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

FEATURES

source

1. 149194

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="7"

/map="7q21-7q22"

/clone="CTB-10G5"

/clone\_lib="CTB-978SK-B"

repeat\_region

1. 1634

/rpt\_family="Li"

misc\_feature

2248. 2388

/note="match to EST BG752883 (NID:g14063536)"

misc\_feature

2248. 2387

/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"

misc\_feature

2248. 2387

/note="match to EST AW579261 (NID:g7254310)"

misc\_feature

2248. 2387

/note="match to EST BG766882 (NID:g14077535)"

misc\_feature

2248. 2387

/note="match to EST BI160365 (NID:g14620366)"

misc\_feature

2248. 2374

/note="similar to Homo sapiens EST BF758865 (NID:g12106765)"

misc\_feature

2248. 2287

/note="match to EST AV686676 (NID:g10288539)"

misc\_feature

2253. 2387

/note="match to EST AU123510 (NID:g10948226)"

misc\_feature

2344. 2387

/note="match to EST AV686676 (NID:g10288539)"

repeat\_region

2696. 3066

/rpt\_family="MaLR"

repeat\_region

3108. 3392

/rpt\_family="Alu"

misc\_feature

3540. 3628



Db	89922	AGGCGGAACAGGACAAACGGGATTAATAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	89863
Qy	1805	CAAGTGACATTTGGAGGCTCTGGAAAAGGGAAGCTGGGCAAAATGAATGCTTAATAGG	1864
Db	89862	CAAGTGACATTTGGAGGCTCTGGAAAAGGGAAGCTGGGCAAAATGAATGCTTAATAGG	89803
Qy	1865	GCTTGTCTTCAGTGGCGTCTACAAGGACACTTTAAAAAAGATTTGTCCAAGTGAAGTAAG	1924
Db	89802	GCTTGTCTTCAGTGGCGTCTACAAGGACACTTTAAAAAAGATTTGTCCAAGTGAAGTAAG	89743
Qy	1925	CCGCCCCCTTCGTTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG	1984
Db	89742	CCGCCCCCTTCGTTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG	89683
Qy	1985	GGACAAAGGCTTTTGTAGTCAGAGCCACTTAACAGATGATCCAGCAGCAGGACTGAGG	2044
Db	89682	GGACAAAGGCTTTTGTAGTCAGAGCCACTTAACAGATGATCCAGCAGCAGGACTGAGG	89623
Qy	2045	TGCTTGGGGCAAGGCCATCCCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACC	2104
Db	89622	TGCTTGGGGCAAGGCCATCCCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACC	89563
Qy	2105	ATTGAGGGCCAGGAAGTGTCTCTGGAACA CTGGTGGGTCTTTTAGTCTTACTCTTC	2164
Db	89562	ATTGAGGGCCAGG-AGGTTGTCTCTGGAACA CTGGTGGGTCTTTTAGTCTTACTCTTC	89504
Qy	2165	TGTCCCGGCAACTGTCTCCAGATCTGTCACTATCTGTAGGGGGTCCNTAAGACGGCA	2224
Db	89503	TGTCCCGGCAACTGTCTCCAGATCTGTCACTA-TCGTAGGGGGTCC- TAAGACGGGCA	89446
Qy	2225	GTCACTAGATACATTTTCCAGCCACTAAGTTATGAATCGGGAGCTTTATTTCTTTTAC	2284
Db	89445	GTCACTAGATAC- TTCTCCCGAGCCACTAAGTTATG- ACTGGGAGCTTTATTTCTTTTAC	89388
Qy	2285	ATGCTTTTCTPAATTAATCTTGAAAGCCCCACTACTCTTTGTTAGGAGAGACAATTTAGCAA	2344
Db	89387	ATGCTTTTCTPAATTAATCTTGAAAGCCCCACTACTCTTTGTTAGGAGAGACAATTTAGCAA	89328
Qy	2345	AAGCAGGGCCNTATACCTGACATAGGAGAGAGAACCCCGTTGTGTNCCCTCG	2404
Db	89327	AAGCAGGGCCNTATATACCTGACATAGGAGAGAGAGAACCCCGTTGTGT- CCCCTG	89269
Qy	2405	CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGACAATATGACAGCCAAA	2464
Db	89268	CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGACAATATGACAG- CAAA	89210
Qy	2465	GAATGCCCGTCTCTTCAAGTTAACTAAAGGATTCACCTTTCCCTTCCCTACCAAGGCA	2524
Db	89209	GAATGCCCGTCTCTTCAAGTTAACTAAAGGATTCACCTTTCCCTTACCAAGGCA	89150
Qy	2525	GTACCCCTCAGACCCAGGCCCAACAGGATTCCAAGATTTGTTAAGACTTAAAGC	2584
Db	89149	GTACCCCTCAGACCCAGGCCCAACAGGATTCCAAGATTTGTTAAGACTTAAAGC	89090
Qy	2585	CCAAGGCTTAGTAAAAACATGCATAACTCCCTGCAGTAATTCCTGATGATTCAGGAGG	2644
Db	89089	CCAAGGCTTAGTAAAAACATGCAGTAATCCCTTCAGTACTCC-----AATTTTAGAG	89037
Qy	2645	CACAGAAACCCAGTGGACAGTGGAGGTTAGTGAAGATCTCAGATATCAATGGAGGC	2704
Db	89036	TACAGAAACCCACACAGACAGTGG- GGTTAGTGAAGATCTCAGATATCAAT- GAGGC	88979
Qy	2705	CGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCTCAATACCAGA	2764
Db	88978	TGTTGTCTCTATAGCCAGCTGTACTAGCCCTTATCTGTCTTTTCCCAATACCAGA	88919
Qy	2765	GGAAGCAGAGTGGTTTACATCTCTGGACCTTMAAGATGCCCTTCTCTGCACTCCCTGTACA	2824
Db	88918	GGAAGCAGAGTGGTTTACATCTCTGGACCTTCAAGATGCCCTTCTCTGCACTCCCTGTACA	88859
Qy	2825	TCCTGACTCTCAATCTTCTTGTGCTTTGAGAGATCTTCAAAACCCACATCTCAACTCAC	2884
Db	88858	TCCTGACTCTCAATCTTCTTGTGCTTTGAGAGATCTTCAAAACCCACATCTCAACTCAC	88799
Qy	2885	CTGGACTTTTACCCAAAGGGTTTCAGGATAGYCCCCATCTATTTTGGCCAGGCATTAGC	2944
Db	88798	CTGGACTTTTACCCAAAGGGTTTCAGGATAGTCCCCATCTATTTTGGCCAGGCATTAGC	88739
Qy	2945	CCAAGACTTGAGYCAATYMTCACTCTGGACACTCTTGTCTCTTCRGTAKGTGATGATTT	3004
Db	88738	CCAAGACTTGAGCAATCTCTATACCTGGACA--CTTGTCTCTCGGTAGGTGATGATTT	88681
Qy	3005	ACTTTTGTGCGCYRTTCAGAAAACCTTGTGCCATCAAGCCACCCAAAGCTCTCTTMAATTT	3064
Db	88680	ACTTTTGTGCGCCATTCAGAAAACCTTGTGCCATCAAGCCACCCAAAGCTCTCTTCAATTT	88621
Qy	3065	CCTCGCVACCTGTGGCTACAWGGTTTCCAAAC SARARGCTCARCTCTGTCTCACAGCAGGT	3124
Db	88620	CCTCGCTACTGTGGCTACATGGTTTCCAAAACCAAGGCTCAACTCTGTCTCACAGCAGGT	88561
Qy	3125	TAAATACTTAGRCTAARATTTATCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYATPCCA	3184
Db	88560	---TACTTAGGCTTAAATATTCAAAGGCACAGGGGCCCTCAGTGAGGAACACATCCA	88505
Qy	3185	GCCTATCTGGCTTATCTCANTCYCAAAAACCTTAAAGCAACTAAGRRRTTCTTCCGCT	3244
Db	88504	GCCTATCTGGCTTATCTCANTCCCAAAAACCTTAAAGCAACTAAGGGGATCTCTTGGCGT	88445
Qy	3245	AAVAGGYTTCTGCCGAATATGATTCCTCCAGGTTTGGCRAAAATAGCCAGGYCAATTAAWATA	3304
Db	88444	AATAGGTTTCTGCCGAATATGATTT-CCCAGGTATGGGAAATAGCCAGGTCAATAATA	88386
Qy	3305	CASTAATTAAGGAACTCAGAAAACCAATACCAATTTATPAAGATGAGAYMCTGAAGYMR	3364
Db	88385	CACATAAAGGAACTCAGAAAACCAATACCAATTTATAGTAGATGACAACTGAAGTAG	88326
Qy	3365	AACTGCTTTCAGGGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAGTGTTAAGYTGCC	3424
Db	88325	AACTGCTTTCAGGGCC-----TAAACCCAGGCCAGTGTTAAGYTGCC	88280
Qy	3425	AACRGGGCAAGACTTTTSTTATYRTCAAGAAAAAACAAGAAVAGCTCTTGGAGTCC	3484
Db	88279	AACRGGGCAAGACTTTTCTTCATATGTCAAG-AAAAACAAGAAATAGCTCTAGGAGTCC	88221
Qy	3485	TTACACAGTCTCAGGAGYAGCTTGCACCCYRTGGCRYACCTGASTAAGAAAYTGA TG	3544
Db	88220	TTACACAGATCCAGGAGTGCAGCTTGCACCTGTGCATACCTGACTAAGGAAATTCATG	88161
Qy	3545	TAGTGGCAAGGGTGTGCTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTCTTACTAT	3604
Db	88160	TAGTGGCAAGGGTGTGACTCTCATTTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTACTAT	88101
Qy	3605	CTGAAGCAGTTAAATAATACAGGAGAGATCTTACTGTGGACATCTCATGAKGTGA	3664
Db	88100	CTGAAGCAGTTAAATAATACAGGAGAGATCTTACTGTGGACATCTCATGATGTA	88041
Qy	3665	AYRGCTATCTCATCTGTAAAGGAGCTTGTGGTGTGCAGACAACTGTTTACTTAAATPTC	3724
Db	88040	ATGCTATCTCATCTGTAAAGGAGCTTGTGGTGTGCAGACAACTGTTTACTTAAATPTC	87981
Qy	3725	AGGCTCTTATCTTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCACTCTTAAACCCAG	3784
Db	87980	AGGCTCTTATCTTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCACTCTTAAACCCAG	87921
Qy	3785	YCNATTTCTTCAGACAAATGAAGAAAGATATAAATACTGTCAACAARTATTTCTC	3844
Db	87920	CCACATTTCTTCAGACAAATGAAGAAAGATATAAATACTGTCAACAARTATTTCTC	87861
Qy	3845	AAAACCTATGCCACTCGAGGGGACTTGTAGAGTTCCTTTGACTGATCCCGACCTTCAAC	3904
Db	87860	AAAACCTATGCCACTCGAGGGGACTTGTAGAGTTCCTTTGACTGATCCCGACCTTCAAC	87802
Qy	3905	TTGTATCTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAGYGGGGTATGCAATG	3964
Db	87801	TTGTATCTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAGYGGGGTATGCAATG	87742



QY	3965	GTCACTGATTAATGGAAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR	4024	Db	86661	AGCNAACCCAGGAACCCCACTCACTATGCCCTGCTCTGTGGCTATAGCCTTAAAGA	86602
Db	87741	GTCACTGATTAATGGAAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTR	87682	QY	5104	ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163
QY	4025	GCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATAGGAGAAAGAAAGGYYAAATATA	4084	Db	86601	ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	86542
Db	87681	GCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATAGGAGAAAGAAAGGYYAAATATA	87622	QY	5164	TTCAATACCAATGACCTTTGCTTGACCCCAAGACAGCAACTTAGTTGCGAGACATCACT	5223
QY	4085	TATACAGACTCTAATATGCTTACCTAGTCTGCTTCCATGCGGCAATATGAGAAATATA	4144	Db	86541	TTCAATACCAATGACCTTTGCTTGACCCCAAGACAGCAACTTAGTTGCGAGACATCACT	86482
Db	87621	TATACAGACTCTAATATGCTTACCTAGTCTGCTTCCATGCGGCAATATGAGAAATATA	87562	QY	5224	CTTTAGCCCAATATCAACAAAGTTCTTAAACATTTCAAGAACTTACCAAGAACTTACCTCTGAGAGAGG	5283
QY	4145	AGGAAATCTCACTTCCGAGGAAACCTATATCMAATCAGGAAGCCATTTAGGAAATATA	4204	Db	86481	CTTTAGCCCAATATCAACAAAGTTCTTAAACATTTCAAGAACTTACCTCTGAGAGAGG	86422
Db	87561	AGGAAATCTCACTTCCGAGGAAACCTATATCMAATCAGGAAGCCATTTAGGAAATATA	87502	QY	5284	GAAGAAGACTTATCCACCCCMWGTGACATGATATAGTCAAGTCCCTTCYCTCTAAATTCCT	5343
QY	4205	TTATGTCGCTGTACAGAAACCTTACAGAGTGTGAGTCTTACACTGCGGGGTCTATCAGAA	4264	Db	86421	GAAGAAGACTTATCCACCCCMWGTGACATGATATAGTCAAGTCCCTTCYCTCTAAATTCCT	86362
Db	87501	TTATGTCGCTGTACAGAAACCTTACAGAGTGTGAGTCTTACACTGCGGGGTCTATCAGAA	87442	QY	5344	CATCCCTAGATACATCTCTGGGAGGACCTTACCAGTCACTTTATTTATACCCCAACTGCGG	5403
QY	4265	GGAAAGGAAAGGGAATASAGGAAATGCAAGCAATATGAGGAAAGGAGCTGCA	4324	Db	86361	CATCCCTAGATACATCTCTGGGAGGACCTTACCAGTCACTTTATTTATACCCCAACTGCGG	86302
Db	87441	GGAAAGGAAAGGGAATASAGGAAATGCAAGCAATATGAGGAAAGGAGCTGCA	87382	QY	5404	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACAATTTAGTCAATCTCTGATGATGTC	5463
QY	4325	AGGAGGACCTCCATTTAGAAATGCTTTAAATCTTCCCTTAGTATAGGGAATCTCCCTTC	4384	Db	86301	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACAATTTAGTCAATCTCTGATGATGTC	86242
Db	87381	AGGAGGACCTCCATTTAGAAATGCTTTAAATCTTCCCTTAGTATAGGGAATCTCCCTTC	87322	QY	5464	CAAAGGAACTGAAATCTCAGGAGACAACTAGTCTTCTGTGAACTCTTAGAGGAT	5523
QY	4385	CGGGAACCAAGCCCACTACTCAGCAGGAGAAACAGAACTGAGGAGCTCAGAGG-CAG	4443	Db	86241	CAAAGGAACTGAAATCTCAGGAGACAACTAGTCTTCTGTGAACTCTTAGAGGAT	86182
Db	87321	CGGGAACCAAGCCCACTACTCAGCAGGAGAAACAGAACTGAGGAGCTCAGAGG-CAG	87262	QY	5524	TGGGCTGCTCTTCAAAACAAACAGGAGGAAAGTAACCTAAATCATAAATCCCCCATG	5583
QY	4444	TTTTCTCCCTCGGAGCGGTAGCCACTGAAAGAGGAAATATCTTTTGCTGCACTAT	4503	Db	86181	TGGGCTGCTCTTCAAAACAAACAGGAGGAAAGTAACCTAAATCATAAATCCCCCATG	86123
Db	87261	TTTTCTCCCTCGGAGCGGTAGCCACTGAAAGAGGAAATATCTTTTGCTGCACTAT	87202	QY	5584	GSCTCTCCCTTATCATATTTTCTCTKATSTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT	5643
QY	4504	CCATGGAAATTAATCTTAAACCCCTTCACTAGGATGATAGGATGATAGGATGATAGGAT	4563	Db	86122	GSCTCTCCCTTATCATATTTTCTCTKATSTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT	86063
Db	87201	CCATGGAAATTAATCTTAAACCCCTTCACTAGGATGATAGGATGATAGGATGATAGGAT	87142	QY	5644	CCCTCTCCATGCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGAAATG	5703
QY	4564	TCARATGGCCAAATCAATTTACTTGGACAGGCTTTTCAAACTATCAAGCAATATAT	4623	Db	86062	CCCTCTCCATGCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGGAGAAATG	86003
Db	87141	TCARATGGCCAAATCAATTTACTTGGACAGGCTTTTCAAACTATCAAGCAATATAT	87082	QY	5704	CAGCGCTCCGGAATATTTGATGCCCTCCATGCTATAGAGTCTTTTSTAAGGAAACCCCAAC	5763
QY	4624	CAGGCTCTGTAAGTGTCCAGAGAAATATCCCTGCTGCTATCGCAAGCTCTTTCAGG	4683	Db	86002	CAGCGCTCCGGAATATTTGATGCCCTCCATGCTATAGAGTCTTTTSTAAGGAAACCCCAAC	85943
Db	87081	CAGGCTCTGTAAGTGTCCAGAGAAATATCCCTGCTGCTATCGCAAGCTCTTTCAGG	87022	QY	5764	TTCACTGCTCCACACCATATGCGCGCAACTGCTATCACTCTGCACTCTTTCATGCAATG	5823
QY	4684	APAAACAAAPAAACAGGCAATTAACCTGGAAPAAACCTGGAACCTGATTTTACCAAGCCC	4743	Db	85942	TTCACTGCTCCACACCATATGCGCGCAACTGCTATCACTCTGCACTCTTTCATGCAATG	85883
Db	87021	APAAACAAAPAAACAGGCAATTAACCTGGAAPAAACCTGGAACCTGATTTTACCAAGCCC	86962	QY	5824	GCAATACTCAATTTATGGAAGGAAATGATTAATCTCTAGTCTCTGAGGACTTGA	5883
QY	4744	AAACCTCAGGATTTTCACTAGTCTGCTGATATATCTTTTCAAGGCTTGGGAGAG	4803	Db	85882	GCAATACTCAATTTATGGAAGGAAATGATTAATCTCTAGTCTCTGAGGACTTGA	85823
Db	86961	AAACCTCAGGATTTTCACTAGTCTGCTGATATATCTTTTCAAGGCTTGGGAGAG	86902	QY	5884	GTCAGTCTGCTGGAATTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTTCAA	5943
QY	4804	GCTTTCCCTGTAGGAGCAAGAGCCCAAGAGGTAATAAGGCACTAGTTCAATGATA	4863	Db	85822	GTCAGTCTGCTGGAATTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTTCAA	85763
Db	86901	GCTTTCCCTGTAGGAGCAAGAGCCCAAGAGGTAATAAGGCACTAGTTCAATGATA	86842	QY	5944	GATCAGGCAAGGAAACATGTAAGAGAGTAATCTCCCACTCACCAGGCTACATGCG	6003
QY	4864	ATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCAGGCG	4923	Db	85762	GATCAGGCAAGGAAACATGTAAGAGAGTAATCTCCCACTCACCAGGCTACATGCG	85703
Db	86841	ATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCAGGCG	86782	QY	6004	ACCTCTAGCCCTCAAGAGGACTAGTCTCTCAAACTACATGAAACCTCTCGTACCCAT	6063
QY	4924	CAGTAACCCAGGAGTATCCAGGCTTATGATATCGATATCACTTACCTGCGCTGA	4983	Db	85702	ACCTCTAGCCCTCAAGAGGACTAGTCTCTCAAACTACATGAAACCTCTCGTACCCAT	85643
Db	86781	CAGTAACCCAGGAGTATCCAGGCTTATGATATCGATATCACTTACCTGCGCTGA	86722	QY	6064	ACTGCTGTGTAAGCTTATTAATACACCTCTCTGAGGCTCCATGAGGCTCTCGGCCAA	6123
QY	4984	AGGCAACAGTCTCAGGAGAGTCCAGGAAATGAAATGAAATCAAGGACATCTAAA	5043	Db	85642	ACTGCTGTGTAAGCTTATTAATACACCTCTCTGAGGCTCCATGAGGCTCTCGGCCAA	85583
Db	86721	AGGCAACAGTCTCAGGAGAGTCCAGGAAATGAAATGAAATCAAGGACATCTAAA	86662	QY	6124	AACCTCTACTACTGTTGATATGCTCCCTGAGTCTCAGGCAATGTTTCAATCCCT	6183
QY	5044	AAGCAACCCAGGAAACCCACTCAGTGGCTGTCTGTTGCTTATAGCCTTAAAGA	5103				

```
Db      85582 AACCTTACTAATCTGTTGGATATGCTCTCCCTGCTGAACCTTCAGGCCATATATGTTTCAATCCCT 85523
Qy      6184  GTACCTGAACAATGGAAACAATCTTCAGCACAGAAATAAACAACCACTTCCTCGTTTATAGTAGGA 6243
Db      85522 GTACCTGAACAAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTCGTTTATAGTAGGA 85463
Qy      6244  CTTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAAACCTTCACCTGTGTAAATTTAGC 6303
Db      85462 CTTCTGTTTCCAAATCTGGAATAAACCCATACCTCAAAACCTTCACCTGTGTAAATTTAGC 85403
Qy      6304  AATACTACATACACAACCAATCCCAATGCATCAGGTGGGTAACTCTCCACACAATA 6363
Db      85402 AATACTACATACACAACCAATCCCAATGCATCAGGTGGGTAACTCTCCACACAATA 85343
Qy      6364  GTCTGCTTACCTCAGGAATAATTTTGTCTGTGCTACCTCAGGCTATCTGTTGTTGAAT 6423
Db      85342 GTCTGCTTACCTCAGGAATAATTTTGTCTGTGCTACCTCAGGCTATCTGTTGTTGAAT 85283
Qy      6424  GGCTTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCYATGRCATCTACACT 6483
Db      85282 GGCTTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCCTATGACCATCTACACT 85223
Qy      6484  GAACAAGATTATACAGTTATGTCTATCTAAGCCCCGCAACAAGAGTACCCATCTTT 6543
Db      85222 GAACAAGATTATACAGTTATGTCTATCTAAGCCCCGCAACAAGAGTACCCATCTTT 85163
Qy      6544  CTTTGTGTTAGGAGCAGGAGTCTAGGTGCATCTAGTACTGTGCATTTGGCGGTATCACA 6603
Db      85162 CTTTGTGTTAGGAGCAGGAGTCTAGGTGCATCTAGTACTGTGCATTTGGCGGTATCACA 85103
Qy      6604  ACCTCTACTCAGTTTCTACTACAAATCTATCTCAAGAACTAAATGGGGCATGGAAACGGGTC 6663
Db      85102 ACCTCTACTCAGTTTCTACTACAAATCTATCTCAAGAACTAAATGGGGCATGGAAACGGGTC 85043
Qy      6664  GCGGACTCCCTGGTCACTTGCAGATCAATTAATCTCCCTAGCAGCAGTAGTCTTCRA 6723
Db      85042 GCGGACTCCCTGGTCACTTGCAGATCAATTAATCTCCCTAGCAGCAGTAGTCTTCRA 84983
Qy      6724  AATCGAAGAGCTTTAGACTCTACCGCTCARAGAGGGGAACCTGTTTATTTTAGGG 6783
Db      84982 AATCGAAGAGCTTTAGACTCTACCGCTCARAGAGGGGAACCTGTTTATTTTAGGG 84923
Qy      6784  GAAGAATGCTGTTTATTTATTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATT 6843
Db      84922 GAAGAATGCTGTTTATTTATTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATT 84863
Qy      6844  CSAGATCGAATACAACTGATAGAGAGAGCTTCGAACACTGCAACCTGGGGCTCTCTC 6903
Db      84862 CGAGATCGAATACAACTGATAGAGAGAGAGCTTCGAACACTGCAACCTGGGGCTCTCTC 84803
Qy      6904  AGCCATGGATGCTGCTGCTGCTCTCCCTCTTAGGACCTCTAGCAGCTATATAATTGCTA 6963
Db      84802 AGCCATGGATGCTGCTGCTGCTCTCCCTCTTAGGACCTCTAGCAGCTATATAATTGCTA 84743
Qy      6964  CTCCTCTTTGACCCCTGTATCTTTTACCTCTCTGTTAACTTTGTCTCTTCAGAAATCGAA 7023
Db      84742 CTCCTCTTTGACCCCTGTATCTTTTAACTCTCTGTTTAACTTTGTCTCTTCAGAAATCGAA 84683
Qy      7024  GCTGTTAAACTTACAAATGGAGCCCAAGATGAGTCCCAAGATAGATCTACCGCAGACCC 7083
Db      84682 GCTGTTAAACTTACAAATGGAGCCCAAGATGAGTCCCAAGATAGATCTACCGCAGACCC 84623
Qy      7084  CTGAGCCGGCTGTAGCCACGATCTGATGTTTAAATGACATCAAGGACCCCTCTCTGAG 7143
Db      84622 CTGAGCCGGCTGTAGCCACGATCTGATGTTTAAATGACATCAAGGACCCCTCTCTGAG 84563
Qy      7144  GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
Db      84562 GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 84503
Qy      7204  SGTGGCCAACTCCCAAGAGCTTATAGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 7263
Db      84502 CGTGGCCAACTCCCAAGAGCTTATAGTTTTCCTGTTGAGATGGGGGACTGAGAGAC 84443
```

```
Qy      7264  AGGACTAGCTGGATTTTCTTAGCTGTAYTAAGAATCCYTAAGCCCTAGTGGGAAGGTGACC 7323
Db      84442 AGGACTAGCTGGATTTTCTTAGCTGTACTAAGAATCCCTAAGCCCTAGTGGGAAGGTGACC 84383
Qy      7324  ACATCCACCTTTTAAACACGGGGCTTGGCAACTTAGTGTACACCTTGACCAATCAGAGAGCTC 7383
Db      84382 ACATCCACCTTTTAAACACGGGGCTTGGCAACTTAGTGTACACCTTGACCAATCAGAGAGCTC 84323
Qy      7384  ACTAAATATGCTTAATTAGGCAAGACAGAGAGTAAAGAAATAGCAATCATCTATTGCTG 7443
Db      84322 ACTAAATATGCTTAATTAGGCAAGACAGAGAGTAAAGAAATAGCAATCATCTATTGCTG 84263
Qy      7444  AGAGCACAGCAGGAGGACCAATGATCGGGATATAAACCAGTGTTCGAGCCGCAACGG 7503
Db      84262 AGAGCACAGCAGGAGGACCAATGATCGGGATATAAACCAGTGTTCGAGCCGCAACGG 84203
Qy      7504  CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563
Db      84202 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 84143
Qy      7564  ATTAATCTTTCGACTGCR 7582
Db      84142 ATTAATCTTTCGACTGCA 84124

RESULT 10
LOCUS   AY101583              10222 bp      DNA      linear      PRI 11-FEB-2004
DEFINITION Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus,
            allele B, complete sequence.
ACCESSION AY101583
VERSION   AY101583.1 GI:37544401
KEYWORDS  .
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10222)
AUTHORS   Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
            Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
            The endogenous retroviral locus ERVWE1 is a bona fide gene involved
            in hominoid placental physiology
            Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL   Direct Submission
PUBMED    14757826
REFERENCE 2 (bases 1 to 10222)
AUTHORS   Mallet,F., Bouton,O. and Oriol,G.
            Retrovirology Department, UMR 2142
            CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
            d'Italie, Lyon 69364 cedex 07, France
FEATURES  Location/Qualifiers
            source
              1..10222
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /isolate="21"
                /isolation_source="PBMC of individual 21"
                /db_xref="taxon:9606"
                /chromosome="7"
                /map="7q21-q22"
                /sex="female"
                /cell_type="PBMC"
                /note="isolated by PCR; ERVWE1 locus, allele B
            endogenous_virus: HERV-W"
              1..780
                join(248..811,2887..9813)
                /note="putative mRNA transcript 1; gag-pol pseudogene"
                join(248..811,7585..9813)
                /product="envelope glycoprotein"
                /note="putative mRNA transcript 2"
                join(248..811,9248..9813)
                /note="putative mRNA transcript 3; no coding region"
                7812..9428
LTR
mRNA
mRNA
mRNA
CDS
```

```

/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68162.1"
/db_xref="GI:37544402"
/translation="MALPYHIFLFTVLLPSFTLTAPPCKMTSSSPYQBFILWRMORP
GNIDATSYRLSLSKGTPTFTTAHMPRNCHYSATLCMHANTHYTGKMINPSCPGGLGV
TWTCTYFTQMGSDGGVQDQAREKHAKEVISQLTRVHGTSYKGLDLSKLRHETLR
HTRLVLFNTTLGLLEHVSQNPNTNCWICLPNRPXVSPVPEOWNFSEINTTNSV
LVGLVSNLEIHTSNLTCVKFSNTTNTSQCIRWVTPPTQVCLPSGIFPVCGRSA
YRCINGSSBQFSLFLVPMWTLTYTBQDLYSVI SKPRNKRPVLPILPVIAGVGLGALG
TGIGGTTSTQFYKLSQELNGDMERVADSLVLOQLNSLAALVQNRALDLITAE
RGTCFLFGECCYVNSQGVTEKVKRIHQDRABELRNTGPNLLSOWMPWTLTP
FLGPLAAILLLGLGPCIFNLNVFSSRIEAVKQLQWPRQSKTKIYRPLDRPASP
RSDVNDIKGTPPEISAQPLLRPNAGSS"
9487. .10222

```

LTR  
ORIGIN

```

Query Match      85.0%; Score 6446.4; DB 9; Length 10222;
Best local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 51; Indels 53; Gaps 27;

QY 606 GATGGGAACGTTCCCGCAGACAAACCGCCCTAAGACGATATTCTGGAATTTGGGA 665
DB 2886 GATGGGAACGTTCCCGCAGACAAACCGCCCTAAGACGATATTCTGGAATTTGGGA 2945

QY 666 MCAATTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTCTCTGACGTGCGGCTG 725
DB 2946 CCAATTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTCTCTGACGTGCGGCTG 3005

QY 726 GCACTCTGAGGAAGTATAAATTATAACCACTCTTACAGTAGACYTCTTTTGTAGAA 785
DB 3006 GCACTCTGAGGAAGTATAAATTATAACCACTCTTACAGTAGACYTCTTTTGTAGAA 3065

QY 786 AAGCAAAATGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAAGAGACAATCAC 845
DB 3066 AAGCAAAATGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAAGAGACAATCAC 3125

QY 846 AATTATGTAAAAAGTGTGATTTATGCCCTACAGAAAGCCCTCAGAGTCTACCTCCCTATC 905
DB 3126 AATTATGTAAAAAGTGTGATTTATGCCCTACAGAAAGCCCTCAGAGTCTACCTCCCTATC 3185

QY 906 CAGCATCTCCCGACTCTCTTCCCACTTAAGAAAGCCCTTCAACCCCAATGGTCCAA 965
DB 3186 CAGCATCTCCCGACTCTCTTCCCACTTAAGAAAGCCCTTCAACCCCAATGGTCCAA 3245

QY 966 AAGGAGATAGACAAAGGTTAAACAGTACCAAGAGTGCCTAATTTCCCAATTTATGA 1025
DB 3246 AAGGAGATAGACAAAGGTTAAACAGTACCAAGAGTGCCTAATTTCCCAATTTATGA 3305

QY 1026 CCCTCCCAAGCAGTGGGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY 1085
DB 3306 CCCTCCCAAGCAGTGGGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY 3364

QY 1086 TCTCCAGACTTAAAGCAATAAATAAAGACTTGTAGTAAATTTCTCAGATAAAYCCTGATGG 1145
DB 3365 TCTCCAGACTTAAAGCAATAAATAAAGACTTGTAGTAAATTTCTCAGATAAAYCCTGATGG 3424

QY 1146 CTATATTGTGTTTTACAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATA 1205
DB 3425 CTATATTGTGTTTTACAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATA 3483

QY 1206 TGTCACTGTCTAAATCAGACACTAACCCCAATCAGAGAGTGCACCAATTAATTCGACGCT 1265
DB 3484 TGTCACTGTCTAAATCAGACACTAACCCCAATCAGAGAGTGCACCAATTAATTCGACGCT 3543

QY 1266 GAGGTTTGGCGATCTCTGTATCTCAGTCAAGTCAATGGATGATGACCAACAGAGG 1325
DB 3544 GAGGTTTGGCGATCTCTGTATCTCAGTCAAGTCAATGGATGATGACCAACAGAGG 3600

QY 1326 AAGANAATGATTTCCCAACAGGCCAGCAGCTTCCAGTCTTASACCTCATTTGGGAC 1385
DB 3601 AAGANAATGATTTCCCAACAGGCCAGCAGCTTCCAGTCTTASACCTCATTTGGGAC 3657

```

```

QY 1386 ACAGAAATCAGTAAACATGGAGATTGGTGTCTGCAGACATTTGCTAACTTGTGTCTASAA 1445
DB 3658 ACAGAAATCAGTAAACATGGAGATTGGTGTCTGCAGACATTTGCTAACTTGTGTGTAGAA 3716

QY 1446 GGACTAAGGAAACCTASGAGAAATCTATGAATTAATCTCAATGATGTCCACCAATACACA 1505
DB 3717 GGACTAAGGAAACCTAAGGAAAG- AAGTCTATGAATTAATCTCAATGATGTCCACCAATACACA 3775

QY 1506 GGGGAGGAGGAGAAATCTCTACTCTGCTCTTCTGAGAGACTAAAGGAGGCAATTGAGGAGC 1565
DB 3776 -GGGAGGAGGAGAAATCTCTACTCTGCTCTTCTGAGAGACTAAAGGAGGCAATTGAGGAGC 3834

QY 1566 GTGCTCTCTGTCTCAGTCTCTCTCTGAAGGCCAACTAACTCTTAAAGCGTAAAGTTTATCA 1625
DB 3835 GTGCTCTCTGTCTCAGTCTCTCTCTGAAGGCCAACTAACTCTTAAAGCGTAAAGTTTATCA 3894

QY 1626 CTCAGTCTGCTGAGACATTTAG- AAAAAAATCTCAAAAGTCTGCCGTAGGCCGAGGACAAA 1684
DB 3895 CTCAGTCTGCTGAGACATTTAGAAAAAATCTCAAAAGTCTGCCGTAGGCCGAGGACAAA 3954

QY 1685 ACTTAGAACCTCTTATGAACCTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGAGC 1744
DB 3955 ACTTAGAACCTCTTATGAACCTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGAGC 4014

QY 1745 AGCGGAAACAGGACAAACGGGATTAATAAAGGCAACCGCTTTAGTCATGACCTCAGG 1804
DB 4015 AGCGGAAACAGGACAAACGGGATTAATAAAGGCAACCGCTTTAGTCATGACCTCAGG 4074

QY 1805 CAAAGTGAATTTGGAGGCTCTGGAAGGAAAGTGGGCAAAATGAAATGCTTAATAG 1864
DB 4075 CAAAGTGAATTTGGAGGCTCTGGAAGGAAAGTGGGCAAAATGAAATGCTTAATAG 4134

QY 1865 GCTTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAGAGATTTCCNAGTAGAGTAAG 1924
DB 4135 GCTTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAGAGATTTCCNAGTAGAGTAAG 4194

QY 1925 CGGCCCTTCTGTCCTCAATGCGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984
DB 4195 CGGCCCTTCTGTCCTCAATGCGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 4254

QY 1985 GGAACAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGCAGGACTGAGGG 2044
DB 4255 GGAACAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGCAGGACTGAGGG 4314

QY 2045 TGCTGGGCAAGCGCCATCCATGCCATCACTCCACAGAGCCCTGGGTTATGCTTGACC 2104
DB 4315 TGCTGGGCAAGCGCCATCCATGCCATCACTCCACAGAGCCCTGGGTTATGCTTGACC 4374

QY 2105 ATTGAGGGCCAGGAGGTTGTCTCTGAGACACTGGTGGGCTCTTTAGTCTTACTCTTC 2164
DB 4375 ATTGAGGGCCAGG- AGGTTGTCTCTGAGACACTGGTGGGCTCTTTAGTCTTACTCTTC 4433

QY 2165 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGACGGGCA 2224
DB 4434 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTA- TCTGAGGGGGTCC- TAAGACGGGCA 4491

QY 2225 GTCACTAGATATCTTTTCCAGCCACTAAGTTATGAAGTGGGAGCTTTATTTCTTTTTCAC 2284
DB 4492 GTCACTAGATATCTTTTCCAGCCACTAAGTTATG- ACTGAGGGAGCTTTATTTCTTTTTCAC 4549

QY 2285 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTTACCTTGTAGGAGAGACATTTCTAGCAA 2344
DB 4550 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTTACCTTGTAGGAGAGACATTTCTAGCAA 4609

QY 2345 AAGCAGGGGCCATTTATACACTGAAACATAGGAGAGGAAACACCCGTTTGTGTNCCCCTG 2404
DB 4610 AAGCAGGGGCCATTTATACACTGAAACATAGGAGAGGAAACACCCGTTTGTGTG- CCCCTG 4668

QY 2405 CTTGAGGAGGAAATTAATCTCTGAGTCTGGGCAACAGAGAGGACATATATGACGAGCCAAA 2464
DB 4669 CTTGAGGAGGAAATTAATCTCTGAGTCTGGGCAACAGAGAGGACATATATGACGAG- CAAA 4727

```



Db	6856	CAGGCGCTGTGAAGTGTGCCAGAGAAAATAATCCCTCCCTTATCGCCAAAGCTCCTTCAGG	6915
Qy	4684	ARAAACAARAACAGGCAATACCTGPARAARACTGGCAACTGATTTTACCCACAAGCCC	4743
Db	6916	AGAACAAAGAACAGGCCATTACCTCGGAGAGACTGGCACTGATTTTACCCACAAGCCC	6975
Qy	4744	AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTABATATCTTTACGGGTGGGCARAG	4803
Db	6976	AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTAGATATCTTTACGGGTGGGCAGAG	7035
Qy	4804	GCCTTCCCTGTAGGACAGAAAAAGCCCAAGAGGTAATAAGGCACCTAGTTCATGAATA	4863
Db	7036	GCCTTCCCTGTAGGACAGAAAAAGCCCAAGAGGTAATAAGGCACCTAGTTCATGAATA	7095
Qy	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	4923
Db	7096	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	7155
Qy	4924	ACAGTAACCCAGGAGTATCCGAGGCTTAGGTATAGGATATACATTTACACTGGGCTGA	4983
Db	7156	ACAGTAACCCAGGAGTATCCGAGGCTTAGGTATAGGATATACATTTACACTGGGCTGA	7215
Qy	4984	AGCCCAAGTCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAAA	5043
Db	7216	AGCCCAAGTCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAAA	7275
Qy	5044	AAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYCTGTGGCTATAGCTTTAAAAGA	5103
Db	7276	AAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYCTGTGGCTATAGCTTTAAAAGA	7335
Qy	5104	ATCTGCAATTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGAAGGCC	5163
Db	7336	ATCTGCAATTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGAAGGCC	7395
Qy	5164	TTCTAATCAATCACTGTGCTTGACCCAGGACGCACTTAGTTGAGACATCACCT	5223
Db	7396	TTCTAATCAATCACTGTGCTTGACCCAGGACGCACTTAGTTGAGACATCACCT	7455
Qy	5224	CCTTAGCCAAATATCAACAGTTCTTTAAACATTAACAGAACTATCCCTGAGAGAGG	5283
Db	7456	CCTTAGCCAAATATCAACAGTTCTTTAAACATTAACAGAACTATCCCTGAGAGAGG	7515
Qy	5284	GAAAGAACTATTCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCCTTAATGCC	5343
Db	7516	GAAAGAACTATTCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCCTTAATGCC	7575
Qy	5344	CATCCCTAGATATCATCTGGGAGGACCTTACCAGTCACTTTATTTACCCCACTGCCG	5403
Db	7576	CATCCCTAGATATCATCTGGGAGGACCTTACCAGTCACTTTATTTACCCCACTGCCG	7635
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATCTGC	5463
Db	7636	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATCTGC	7695
Qy	5464	CAAAGAACTGAAATTCAGGAGACAAGCTAGCTATTCCTGTGAACCTCTAGAGGAT	5523
Db	7696	CAAAGAACTGAAATTCAGGAGACAAGCTAGCTATTCCTGTGAACCTCTAGAGGAT	7755
Qy	5524	TGGCCCTGCTTCAACACACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG	5583
Db	7756	TGGCCCTGCTTCAACACACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG	7814
Qy	5584	GSCTCCCTTATCATATTTTCTKTATSTGTTSTTTTACCTSTTTTCACTCTCACTGCA	5643
Db	7815	GSCTCCCTTATCATATTTTCTKTATSTGTTSTTTTACCTSTTTTCACTCTCACTGCA	7874
Qy	5644	CCCCCTCCATGCGCTGTATGACACAGTACCTCCCTTACCCAGAGTTTCTATGAGGATG	5703
Db	7875	CCCCCTCCATGCGCTGTATGACACAGTACCTCCCTTACCCAGAGTTTCTATGAGGATG	7934
Qy	5704	CAGCGTCCCGGAAATATGATGCCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAAC	5763
Db	7935	CAGCGTCCCGGAAATATGATGCCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAAC	7994
Qy	5764	TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTCGCACTCTTTGCATGCAT	5823
Db	7995	TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTCGCACTCTTTGCATGCAT	8054
Qy	5824	GCAAAATCTCATTTATTTGGAAGGAAATAATGAATTAATCTTAGTTGTCTCGGAGGACTTGA	5883
Db	8055	GCAAAATCTCATTTATTTGGAAGGAAATAATGAATTAATCTTAGTTGTCTCGGAGGACTTGA	8114
Qy	5884	GTCACTGTCTGTGGACTTACTTACCCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAA	5943
Db	8115	GTCACTGTCTGTGGACTTACTTACCCCAAACTGGTATGTCTGATGGGGGTGGAGTTCAA	8174
Qy	5944	GATCAGCAGAGAGAAAAACATGTAAAAAGAAATTAATCTCCCAACTCACCCGGGTACATGGC	6003
Db	8175	GATCAGCAGAGAGAAAAACATGTAAAAAGAAATTAATCTCCCAACTCACCCGGGTACATGGC	8234
Qy	6004	ACCTTAGCCCTTAACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCAT	6063
Db	8235	ACCTTAGCCCTTAACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCAT	8294
Qy	6064	ACTCGCCTCGTAAGCCCTATTTAATACCAACCCCTCACTGGGCTCCATGAGTCTCGGCCAA	6123
Db	8295	ACTCGCCTCGTAAGCCCTATTTAATACCAACCCCTCACTGGGCTCCATGAGTCTCGGCCAA	8354
Qy	6124	AACCTTACTAACTGTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	8355	AACCTTACTAACTGTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	8414
Qy	6184	GTACTCGAAATCAATGGAACAACTTCAGCAGAGAAATAAAACCACTTCCGTTTTAGTAGGA	6243
Db	8415	GTACTCGAAATCAATGGAACAACTTCAGCAGAGAAATAAAACCACTTCCGTTTTAGTAGGA	8474
Qy	6244	CCTCTGTCTCCAAATSTGGAAATACCCATACCTCAAACTCACTGCTGTGTAATTTAGC	6303
Db	8475	CCTCTGTCTCCAAATSTGGAAATACCCATACCTCAAACTCACTGCTGTGTAATTTAGC	8534
Qy	6304	AATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAATA	6363
Db	8535	AATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAATA	8594
Qy	6364	GTCTGCTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6423
Db	8595	GTCTGCTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8654
Qy	6424	GGCTCTTCAGAAATCTATGTGTCTCTCAATCTTAGTGTGTGTGTGTGTGTGTGTGTGT	6483
Db	8655	GGCTCTTCAGAAATCTATGTGTCTCTCAATCTTAGTGTGTGTGTGTGTGTGTGTGTGT	8714
Qy	6484	GAAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAAGAGGTACCCATTTCT	6543
Db	8715	GAAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAAGAGGTACCCATTTCT	8774
Qy	6544	CCTTTTGTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGGCAATCGCGGTATCACA	6603
Db	8775	CCTTTTGTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGGCAATCGCGGTATCACA	8834
Qy	6604	AGCTCTACTCAGTCTTCTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTC	6663
Db	8835	AGCTCTACTCAGTCTTCTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGGTC	8894
Qy	6664	GCCGACTCCCTGGTCACTTTCAGAGATCAATTTAACTCCCTAGCAGCAGTAGTCTTCTCA	6723
Db	8895	GCCGACTCCCTGGTCACTTTCAGAGATCAATTTAACTCCCTAGCAGCAGTAGTCTTCTCA	8954
Qy	6724	AATCGAAGAGCTTTAGACTGTCTAACCCGCTGAPAGAGGGGAACTCTTTATTTTAGGG	6783
Db	8955	AATCGAAGAGCTTTAGACTGTCTAACCCGCTGAPAGAGGGGAACTCTTTATTTTAGGG	9014
Qy	6784	GAAAGATGCTGTATTTATTTTAAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAGAAAT	6843
Db	9015	GAAAGATGCTGTATTTATTTTAAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAGAAAT	9074





QY 966 AAGGAGATAGACAAAAGGGTAAAACAGTGCAACCAAGAGTGCCAAATATTTCCCAATATGA 1025  
DB |||||  
QY 3246 AAGGAGATAGACAAAAGGGTAAAACAGTGCAACCAAGAGTGCCAAATATTTCCCAATATGA 3305  
DB |||||  
QY 1026 CCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCCAGAGTGCAATGTCYTTTTYY 1085  
DB |||||  
QY 3306 CCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCCAGAGTGCAATGTCYTTTTYY 3364  
DB |||||  
QY 1086 TCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCCAGAGTGCAATGTCYTTTTYY 1145  
DB |||||  
QY 3365 TCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCCAGAGTGCAATGTCYTTTTYY 3424  
DB |||||  
QY 1146 CTATATGRTGTTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGAGAGATATATA 1205  
DB |||||  
QY 3425 CTATATGRTGTTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGAGAGATATA-A 3483  
DB |||||  
QY 1206 TGTCACTGCTTAATCAGACACTAACCCCAATAGAGAAAGTGCCACATAATCTGACGCT 1265  
DB |||||  
QY 3484 TGTCACTGCTTAATCAGACACTAACCCCAATAGAGAAAGTGCCACATAATCTGACGCT 3543  
DB |||||  
QY 1266 GAGGRTTGGCGATCTCTGTATCTCAGTCAGGCTCAATGGATANGATGACACAGAG 1325  
DB |||||  
QY 3544 GAGGRTTGGCGATCTCTGTATCTCAGTCAGGCTCAATGGATANGATGACACAGAG 3600  
DB |||||  
QY 1326 AAGGANAATGATTTCCCAAGCCAGCAGCAGTCTCCAGTCTPASACCTCTCATTTGGGAC 1385  
DB |||||  
QY 3601 AAGGANAATGATTTCCCAAGCCAGCAGCAGTCTCCAGTCTPASACCTCTCATTTGGG 3657  
DB |||||  
QY 1386 ACAGAAATCAGTAAATGAGGAGATTTGGTGTCAGACATTTGCTAACTTTGCTGCTASAA 1445  
DB |||||  
QY 3658 ACAGAAATCAGTAAATGAGGAGATTTGGTGTCAGACATTTGCTAACTTTGCTGCTAGAA 3716  
DB |||||  
QY 1446 GCACTAAGGAAACTAGAGAAATCTAGTAATCTCAATGATCTCAATGATCTCAATGATCTCA 1505  
DB |||||  
QY 3717 GCACTAAGGAAACTAGAGAAATCTAGTAATCTCAATGATCTCAATGATCTCAATGATCTCA 3775  
DB |||||  
QY 1506 GGGGAGGGAAGAAATCTACTGCTCTCTGAGAGACTAAGGGAGGCAATTTAGGAGC 1565  
DB |||||  
QY 3776 -GGGAGGGAAGAAATCTACTGCTCTCTGAGAGACTAAGGGAGGCAATTTAGGAGC 3834  
DB |||||  
QY 1566 GTGCTCTCTGTCACTGATCTCTGAAAGGCAACTAATCTTAAGCGTAAATGTTATCA 1625  
DB |||||  
QY 3835 GTGCTCTCTGTCACTGATCTCTGAAAGGCAACTAATCTTAAGCGTAAATGTTATCA 3894  
DB |||||  
QY 1626 CTCAGTCAGTCGACACATTTAG-AAAAAATCTCAAGGCTGCGCTAGGCCGAGCAAA 1684  
DB |||||  
QY 3895 CTCAGTCAGTCGACACATTTAGAAAAAATCTCAAGGCTGCGCTAGGCCGAGCAAA 3954  
DB |||||  
QY 1685 ACTTAGAAACCCATTTAGAACTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 1744  
DB |||||  
QY 3955 ACTTAGAAACCCATTTAGAACTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 4014  
DB |||||  
QY 1745 AGGCGGAACAGGACAAAAGGATTTAAAAAAGGCAACCGCTTTAGTCATGACCTCAGG 1804  
DB |||||  
QY 4015 AGGCGGAACAGGACAAAAGGATTTAAAAAAGGCAACCGCTTTAGTCATGACCTCAGG 4074  
DB |||||  
QY 1805 CAACTGGAGCTTTGGAGGCTCTGGAAGAGGAAAGCTGGCAAAATTTGAATGCTTAATAGG 1864  
DB |||||  
QY 4075 CAACTGGAGCTTTGGAGGCTCTGGAAGAGGAAAGCTGGCAAAATTTGAATGCTTAATAGG 4134  
DB |||||  
QY 1865 GCTTGCTTCCAGTGGGCTCTACAGGACACTTTTAAAAAAGATTTGCTCAAGTAGAGTAAG 1924  
DB |||||  
QY 4135 GCTTGCTTCCAGTGGGCTCTACAGGACACTTTTAAAAAAGATTTGCTCAAGTAGAGTAAG 4194  
DB |||||  
QY 1925 CGGCCCTTCTGTCATGCCCCCTTTTCAAGGGAATCTAGGAGGCCCACTGGCCCCAGG 1984  
DB |||||  
QY 4195 CGGCCCTTCTGTCATGCCCCCTTTTCAAGGGAATCTAGGAGGCCCACTGGCCCCAGG 4254  
DB |||||  
QY 1985 GGAACAAGGCTTTTGTAGTCAGAGCCACTAACCGATGATCCAGAGCAGGACTGAGGG 2044  
DB |||||  
QY 4255 GGAACAAGGCTTCTGTAGTCAGAGCCACTAACCGATGATCCAGAGCAGGACTGAGGG 4314  
DB |||||

QY 2045 TGCCTGGGCGAAGCGCCATCCCATGCAATCACCTCAAGAGCCCTGGGTATGCTTGACC 2104  
DB |||||  
QY 4315 TGCCTGGGCGAAGCGCCATCCCATGCAATCACCTCAAGAGCCCTGGGTATGCTTGACC 4374  
DB |||||  
QY 2105 ATTGAGGGCGAGGAAGGTTGTCTCTCGGACACTGCTGGTCTTCTTACTTCTTCTTCTT 2164  
DB |||||  
QY 4375 ATTGAGGGCGAGG-AGGTTGTCTCTCGGACACTGCTGGTCTTCTTACTTCTTCTT 4433  
DB |||||  
QY 2165 TGTCCCGGCAACTGCTGCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCA 2224  
DB |||||  
QY 4434 TGTCCCGGCAACTGCTGCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 4491  
DB |||||  
QY 2225 GTCACCTAGATACATTTTCCAGGCCACTAAGTTATGAATCTGGGGAGCTTTATTCTTTT 2284  
DB |||||  
QY 4492 GTCACCTAGATAC-TTCTCCCGGCACTAAGTTATG-ACTGGGGAGCTTTATTCTTTT 4549  
DB |||||  
QY 2285 ATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA 2344  
DB |||||  
QY 4550 ATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA 4609  
DB |||||  
QY 2345 AAGCAGGGGCGCAATTAATACCTGAAATAGGAGAGGAAACACCCGTTTGTGTNCCCTG 2404  
DB |||||  
QY 4610 AAGCAGGGGCGCAATTAATACCTGAAATAGGAGAGGAAACACCCGTTTGTGT-CCCC 4668  
DB |||||  
QY 2405 CTTGAGGAGGAATTAATTCCTGAACTCTGGGCAACAGAGGCAATATGGAAGGCAAA 2464  
DB |||||  
QY 4669 CTTGAGGAGGAATTAATTCCTGAACTCTGGGCAACAGAGGCAATATGGAAGGCA-CAAA 4727  
DB |||||  
QY 2465 GAATGCCGCTCTGTTTCAAGTTTAAAGATTTAAAGATTTCAAGATTTGTTAAGGACTTAAAGC 2524  
DB |||||  
QY 4728 GAATGCCGCTCTGTTTCAAGTTTAAAGATTTAAAGATTTCAAGTCTTCTTCTTCAAGGCA 4787  
DB |||||  
QY 2525 GTACCCCTCTAGACCCCAAGGCGCAACAAAGGATTTCAAAAGATTTGTTAAGGACTTAAAGC 2584  
DB |||||  
QY 4788 GTACCCCTCTAGACCCCAAGGCGCAACAAAGGACTTCAAAAGATTTGTTAAGGACTTAAAGC 4847  
DB |||||  
QY 2585 CCAAGGCTTAGTAAACCATGCAATCTCCCTGAGTAACTCCCTGAGTAACTCCCTGAGGAGG 2644  
DB |||||  
QY 4848 CCAAGGCTTAGTAAACCATGCAATCTCCCTGAGTAACTCCCTGAGTAACTCCCTGAGGAG 4900  
DB |||||  
QY 2645 CACGAAACCCAGCTGGAAGGCTTAGTGAAGATTTAGGAGATTTCAAGATTTCAATGGAGC 2704  
DB |||||  
QY 4901 TACGAAACCCCAACAGACAGTGGA-AGTTAGTGAAGATTTCAAGATTTCAAT-AGGC 4958  
DB |||||  
QY 2705 GGTGCTCTTTTATACCCAGCTGACCTAGCCCTTATCTGTGMYTTTCCCAATACCAGA 2764  
DB |||||  
QY 4959 GTTGTCTCTTATAGCCAGCTGACCTAGCCCTTATCTGTGCTTTCCCAATACCAGA 5018  
DB |||||  
QY 2765 GGAAGCAGGTTGTTTACATCTGGAAGGCTTGAAGATGCTTCTGATCCCTGTACA 2824  
DB |||||  
QY 5019 GGAAGCAGGTTGTTTACATCTGGAAGGCTTGAAGATGCTTCTGATCCCTGTACA 5078  
DB |||||  
QY 2825 TCCTGACTCTCAATTTCTTGTGCTTTGAAGATTAATCAACCCCAAGATTTCAACTCAC 2884  
DB |||||  
QY 5079 TCCTGACTCTCAATTTCTTGTGCTTTGAAGATTAATCAACCCCAAGATTTCAACTCAC 5138  
DB |||||  
QY 2885 CTGGAATTTTATACCCAGGCTTCAAGGATGAGTCCCATCTATTGTCAGGCAATAGC 2944  
DB |||||  
QY 5139 CTGGAATTTTATACCCAGGCTTCAAGGATGAGTCCCATCTATTGTCAGGCAATAGC 5198  
DB |||||  
QY 2945 CCAAGACTGAGYCARPTMTACATCTGGAACCTTGTCTTCTGATGAGTGAATTT 3004  
DB |||||  
QY 5199 CCAAGACTGAGYCARPTMTACATCTGGAACCTTGTCTTCTGATGAGTGAATTT 5256  
DB |||||  
QY 3005 ACTTTTGGCTGCTTCAAGAACTTTGTCATCAAGGCAACCCCAAGGCTTCTTMAATTT 3064  
DB |||||  
QY 5257 ACTTTTGGCTGCTTCAAGAACTTTGTCATCAAGGCAACCCCAAGGCTTCTTCAATTT 5316  
DB |||||  
QY 3065 CTTGCGYACTGTTGGCTACAGGTTTCCAAACCSARAGCTCACTCTGCTCAGACAGGT 3124  
DB |||||  
QY 5317 CTTGCGYACTGTTGGCTACAGGTTTCCAAACCSARAGCTCACTCTGCTCAGACAGGT 5376  
DB |||||  
QY 3125 TAAATACTTAGGCTAATATTTCAAGGCAACCGGCGCTCAGTGGGAAATYATCCA 3184  
DB |||||

Db	5377	:     :
----	------	---



```
Db 9735 CAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTGTGTTTTCATGCTATTTCACTCT 9794
QY 7564 ATTAATACTTGCARCTGCR 7582
Db 9795 ATTAATACTTGCARCTGCA 9813

RESULT 12
BD221808
LOCUS BD221808 10499 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION BD221808
VERSION BD221808.1 GI:33031578
KEYWORDS JP 2002518051-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A 3 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/3
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09, A01K67/027, A61K31/711, A61K48/00, A61P21/00, PC
A61P25/00.
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..10499
/organism="Homo sapiens (human)".
FEATURES
source 1..10499
/organisms="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 84.9%; Score 6436; DB 6; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;
QY 606 GATGGGAAACGTTCCCGCAGACAAACAAACGCCCCCTAAGCGTATTTCTGGARAAATTGGGA 665
Db 2953 GATGGGAAACGTTCCCGCAGACAAACAAACGCCCCCTAAGCGTATTTCTGGARAAATTGGGA 3012
QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTCTGCGAGTCCGCGCTG 725
Db 3013 CCAATTTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTCTGCGAGTCCGCGCTG 3072
QY 726 GCACCTCTGAGGGAAGTATAAATATTAACACCATCTTTACAGCTAGACYTCTTTTGTAGAA 785
Db 3073 GCACCTCTGAGGGAAGTATAAATATTAACACCATCTTTACAGCTAGACCTCTTTGTAGAA 3132
QY 786 AAGGCAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCTTAAGAGACAACTCAC 845
Db 3133 AAGGCAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCTTAAGAGACAACTCAC 3192
QY 846 AATTATGTAAGAGTGATTTATGCCCCCTACAGAGAGCCCTTACAGAGTCTACCTCCCTATC 905
Db 3193 AATTATGTAAGAGTGATTTATGCCCCCTACAGAGAGCCCTTACAGAGTCTACCTCCCTATC 3252
QY 906 CCAGCATCCCGGACTCTCTCCCAATTAAGAGACCCCTTCAACCCCAATGGTCCAA 965
```

```
Db 3253 CCAGCATCCCGGACTCTCTCCCAATTAAGAGACCCCTTCAACCCCAATGGTCCAA 3312
QY 966 AAGGAGATACACAAAGGGTAAACAGTGAAACCAAGAGTGCCCAATATTTCCCAATTTGA 1025
Db 3313 AAGGAGATACACAAAGGGTAAACAGTGAAACCAAGAGTGCCCAATATTTCCCAATTTGA 3372
QY 1026 CCCCTCCCAAGCAGTGGGAGGAAAGAAATTCGGGCCAGCAGAGTGATGTCGCTTTTYY 1085
Db 3373 CCCCT-CCAAGCAGTGGGAGGAAAGAAATTCGGGCCAGCAGAGTGATGTCGCTTTTTC 3431
QY 1086 TCTCCAGACTTAAAGCAAAATAAAACAGACTTTAGGTAAATTTCTCAGATAAYCCTGATGG 1145
Db 3432 TCTCCAGACTTAAAGCAAAATAAAACAGACTTTAGGTAAATTTCTCAGATAACCCCTGATGG 3491
QY 1146 CTATATTTGTTTTCACAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATATA 1205
Db 3492 CTATATTTGATTTTCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATA-A 3550
QY 1206 TGTCACTGTCTAAATCAGACACTAACCCCAATCAGAGAGTGGCCACCATTAACCTGCAGCCT 1265
Db 3551 TGTCACTGTCTAAATCAGACACTAACCCCAATCAGAGAGTGGCCACCATTAACCTGCAGCCT 3610
QY 1266 GAGGTGTCGATCTCTGATATCTCAGTCAGTCAATGGATANGGATGACAAAGAGAGG 1325
Db 3611 GAGAGTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACACAG-AGG 3667
QY 1326 AAGGANAATGATTTCCCAAGGCCAGCAGCAGTCCCACTCTASACCTCATTTGGGAC 1385
Db 3668 AAGAGAAATGATTTCCCAAGGCCAGCAGCAGTCCCACTCTAGACCTCATTTGGG-- 3724
QY 1386 ACAGAAATCAGTAACATGGAGATTTGGTGTCTGACAGACATTTGCTAACTTGTGTCTASAA 1445
Db 3725 ACAGAAATCAGTAACAT-GGAGATTTGGTGTCTGACAGACATTTGCTAACTTGTGTCTAGAA 3783
QY 1446 GGAATAAGGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCAATAACACA 1505
Db 3784 GGAATAAGGAAACTAGGAAG-AAGTCTATGAATTAATCAATGATGTCACCAATAACACA 3842
QY 1506 GGGGAGGAGGAGAAATCTTACTGCTTCTGAGAGACTAAGGGGAGGCTTGGAGAGC 1565
Db 3843 -GGGAGGAGGAGGAGAAATCTTACTGCTTCTGAGAGACTAAGGGGAGGCTTGGAGAGC 3901
QY 1566 GTGCTCTCTGTCTCACTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAAGTTTATCA 1625
Db 3902 GTGCTCTCTGTCTCACTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAAGTTTATCA 3961
QY 1626 CTGAGTCAGTCGACATTTAG--AAAAAATTTCAAAAGTCTGCGTATAGGCCCGGAGCAAA 1684
Db 3962 CTCAGTCAGTCGACATTTAGAAAAAAATTTCAAAAGTCTGCGTATAGGCCCGGAGCAAA 4021
QY 1685 ACTTAGAAACCTTTGAACTTGGCAACTTGGCTTTTATATATAGATCAGAGAGAGC 1744
Db 4022 ACTTAGAAACCTTTGAACTTGGCAACTTGGCTTTTATATATAGATCAGAGAGAGC 4081
QY 1745 AGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCAACCCCTTTAGTCATGACCCCTCAGG 1804
Db 4082 AGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCAACCCCTTTAGTCATGACCCCTCAGG 4141
QY 1805 CAAGTGGACTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGG 1864
Db 4142 CAAGTGGACTTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAAAATTTGAATGCTTAATAGG 4201
QY 1865 GCTTGTCTTCAGTGGCGTCTACAAGGACACTTTAAAAAAGATTTGTCAGATGAGAGTAAG 1924
Db 4202 GCTTGTCTTCAGTGGCGTCTACAAGGACACTTTAAAAAAGATTTGTCAGATGAGAGTAAG 4261
QY 1925 CCGCCCTCTCTGCTCCATGCCCCCTTTATTTCAAGGGAATCAGTCGAGGCGCCCTATGCCCCAGG 1984
Db 4262 CCGCCCTCTCTGCTCCATGCCCCCTTTATTTCAAGGGAATCAGTCGAGGCGCCCTATGCCCCAGG 4321
QY 1985 GGAACAAAGGTCTTTTGTAGTCAGAGGACCTTAAACAGATATCCAGCAGCAGGACTGAGGG 2044
```

Db	4322	GGACAAGGTCCTCTGAGTCAGAAAGCCACTAACACGATGATCCACGACGAGCACTGAGGG	4381	Qy	3125	TAAATCTTAGGRCATTAATTCCTCAAGGCAACGAGCCCTCAGTCAGGAAATATATCCA	3184
Qy	2045	TGCTGGGGCAAGGCCATCCATGCCATACCCCTCACAGAGCCCTGGGTATGCTGACC	2104	Db	5444	-----TACTTAGGGCTAAATATATCCAAAGGCAACAGGGCCCTCAGTCAGGAAACATCCA	5499
Db	4382	TGCTGGGGCAAGGCCATCCATGCCATACCCCTCACAGAGCCCTGGGTATGCTGACC	4441	Qy	3185	GCCTATCTGGCTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTCTGGCRT	3244
Qy	2105	ATTGAGGCCAGGAAGGTTGTCTCTGGACACTGGTGGGTTCTTTAGTCTTACTCTTC	2164	Db	5500	GCCTATCTGGCTTATCTCTCATCCCAAAACCCCTAAGCAACTAAGGGGATCTCTGGCGT	5559
Db	4442	ATTGAGGCCAGG-AGGTTGTCTCTGGACACTGGTGGGTTCTTTAGTCTTACTCTTC	4500	Qy	3245	AAVAGGTTCTGGCGAAATGGAATCCCGAGGTGGGCRAAATAGCCAGGYCATTAATA	3304
Qy	2165	TGTCGGGACAACTGCTCTCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGCA	2224	Db	5560	AATAGGTTCTGGCGAAATGGAAT-CCAGGTATGGCAATAGCCAGGTCTAATAATA	5618
Db	4501	TGTCGGGACAACTGCTCTCAGATCTGTCACTA-TCTGAGGGGTTCC-TAAGACGGCA	4558	Qy	3305	CAGTAATTAAGGAACCTCAGAAAGCCAAATACCCATTTARTAAAGATGAYAMCTCAAGYMR	3364
Qy	2225	GTCACTAGATCTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATCTTTTCAC	2284	Db	5619	CATTAATTAAGGAACCTCAGAAAGCCAAATACCCATTTAGTAAGATGGAACACTGAAGTAG	5678
Db	4559	GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGAGCTTTATCTTTTCAC	4616	Qy	3365	AAGTGGCTTTCCAGGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCAGTGTTAAGYTTGCC	3424
Qy	2285	ATGCTTTTCTAATATGCTTGAAGGCCCACTACCTTTGTTAGGGAGAGACATTTAGCAA	2344	Db	5679	AAGTGGCTTTCCAGGGCC-----TAAACCAAGCCCAAGTGTTAAGYTTGCC	5724
Db	4617	ATGCTTTTCTAATATGCTTGAAGGCCCACTACCTTTGTTAGGGAGAGACATTTAGCAA	4676	Qy	3425	AACRGGCAAGACTTTTSTTTTATATATATATATATATATATATATATATATATATATAT	3484
Qy	2345	AAGCAGGGCCATTTATACACTGAACTAGAGAGAGAAACACCCGTTTGTGTTGTTGTTG	2404	Db	5725	AACAGGGCAAGACTTTTCTTATATATATATATATATATATATATATATATATATATAT	5783
Db	4677	AAGCAGGGCCATTTATACACTGAACTAGAGAGAGAAACACCCGTTTGTGTTGTTGTTG	4735	Qy	3485	TTACACAGTCCBAGGAGGAGCTTGCACACCTTGGCRVACCTGASTAAGGAAATTCATG	3544
Qy	2405	CTTGAGGAAGAAATTAATCTGAACTCTGGGCAACAGAGACAAATATGACGAGCAAA	2464	Db	5784	TTACACAGATCCGAGGGATGAGCTTGCACCTTGGCATACCTGACTAAGGAAATTCATG	5843
Db	4736	CTTGAGGAAGAAATTAATCTGAACTCTGGGCAACAGAGACAAATATGACGAG-CAAA	4794	Qy	3545	TAGTGGCAAGGCTTGRCTTCATTTTATYGGGTAGTGGTGGCAGTAGCAGTGYKTATAT	3604
Qy	2465	GAATGCCGCTCTCTTCAAGTTAAACTAAAGGATTCACATTCCTTCCCTACCAAGGCCA	2524	Db	5844	TAGTGGCAAGGCTTGRCTTCATTTTATYGGGTAGTGGTGGCAGTAGCAGTGYKTATAT	5903
Db	4795	GAATGCCGCTCTCTTCAAGTTAAACTAAAGGATTCACATTCCTTCCCTACCAAGGCCA	4854	Qy	3605	CTGAAGCAGTTAAATAATAATAACAGGAGAGATCTTACTGTGTGGACATCTCATGATGTA	3664
Qy	2525	GTACCCCTCAGACCCCAAGGCCCAACAGGATTCACAAAGATTTGTTAAGGACTTAAAGC	2584	Db	5904	CTGAAGCAGTTAAATAATAATAACAGGAGAGATCTTACTGTGTGGACATCTCATGATGTA	5963
Db	4855	GTACCCCTCAGACCCCAAGGCCCAACAGGATTCACAAAGATTTGTTAAGGACTTAAAGC	4914	Qy	3665	AYGACATCTACTGCTTAAAGGAGACTTGTGGCTGTGACAACTCTTAAATATTC	3724
Qy	2585	CCAAGGCTAGTAAACCAATGCACTACCTCGAGTAATTCCTGAGTGGATTCAGAGG	2644	Db	5964	AYGACATCTACTGCTTAAAGGAGACTTGTGGCTGTGACAACTCTTAAATATTC	6023
Db	4915	CCAAGGCTAGTAAACCAATGCACTACCTCGAGTAATTCCTGAGTGGATTCAGAGG	4967	Qy	3725	AGGCTCTATTACTTGAAGGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3784
Qy	2645	CACAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATCTCAGGATTTCAATGGAGGC	2704	Db	6024	AGGCTCTATTACTTGAAGGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6083
Db	4968	TACAGAAACCCACAGACAGTGA-GGTAGTCAAGATCTCAGGATTTCAAT-AGGCG	5025	Qy	3785	YCNATTTCTTCCAGCAATATGAAGAAAGATAPAAATATCTCTCAACAARTATTTCTC	3844
Qy	2705	CGTTGCTTTTATACCCAGTGTACCTAGCCCTTATCTGTGMYTTCCCAATATACAGA	2764	Db	6084	CCACATTTCTTCCAGCAATATGAAGAAAGATAPAAATATCTCTCAACAARTATTTCTC	6143
Db	5026	TGTTGTTCTCTATAGCCAGCTGTACCTAGCCCTTATCTGCTTTCCCAATATCCAGA	5085	Qy	3845	AAACCTATGCCACTCCAGGGGACCTTGTAGAGTTCCYTTGACTGATCCYGACCTCAAC	3904
Qy	2765	GGAAGCAGTGGTTTACASTCCCTGGACCTTMAAGTGCCTCTTCTGCTATCCCTGTACA	2824	Db	6144	AAACCTATGCCACTCCAGGGGACCTTGTAGAGTTCCYTTGACTGATCCYGACCTCAAC	6202
Db	5086	GGAAGCAGTGGTTTACASTCCCTGGACCTTMAAGTGCCTCTTCTGCTATCCCTGTACA	5145	Qy	3905	TTGTATCTATGATGGAAGTTCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG	3964
Qy	2825	TCCTGACTCTCAATTTCTGTTTGGCTTTGAAGTACTTCAAAACCCACATCTCAACTCAC	2884	Db	6203	TTGTATCTATGATGGAAGTTCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG	6262
Db	5146	TCCTGACTCTCAATTTCTGTTTGGCTTTGAAGTACTTCAAAACCCACATCTCAACTCAC	5205	Qy	3965	GTCACTGATTAATGGAATATVTTGAAAGTATCCCTCATCTCCAGGAACTAGTGTCTYAGCTR	4024
Qy	2885	CTGGAATTTTATCCCAAGGTTTCAGGATAGYCCCATCTATTTTGGCCAGGATTTAGC	2944	Db	6263	GTCACTGATTAATGGAATATVTTGAAAGTATCCCTCATCTCCAGGAACTAGTGTCTYAGCTR	6322
Db	5206	CTGGAATTTTATCCCAAGGTTTCAGGATAGYCCCATCTATTTTGGCCAGGATTTAGC	5265	Qy	4025	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTTAGGAGAGAAAGGGYAAATATA	4084
Qy	2945	CCAAGACTCAGYCARVMTCACTACCTGGACCTCTGCTGCTCTGCTGCTGCTGCTGCTGCT	3004	Db	6323	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTTAGGAGAGAAAGGGYAAATATA	6382
Db	5266	CCAAGACTCAGYCARVMTCACTACCTGGACCTCTGCTGCTCTGCTGCTGCTGCTGCTGCT	5323	Qy	4085	TATACAGACTCTCAATATCTYACCTAGTCTNTCCATGCCCATGMRGCAATATATGABAGAA	4144
Qy	3005	ACTTTTGGCGCCCATTTAGAAACCTTGTGGCATCAAGCCACCAAGCCCTCTTMAATTT	3064	Db	6383	TATACAGACTCTCAATATCTYACCTAGTCTNTCCATGCCCATGMRGCAATATATGABAGAA	6442
Db	5324	ACTTTTGGCGCCCATTTAGAAACCTTGTGGCATCAAGCCACCAAGCCCTCTTMAATTT	5383	Qy	4145	AGGGAATTCCTAATCTTCYAGGAGCAACCTATCAMACATCAGGAAGCCATTAGGAATTA	4204
Qy	3065	CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3124	Db	6443	AGGGAATTCCTAATCTTCYAGGAGCAACCTATCAMACATCAGGAAGCCATTAGGAATTA	6502
Db	5384	CCTCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5443				





```
Db 8662 GTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTAGCTCAGCTATCGTTGTGAAT 8721
Qy 6424 GGCTCTTCAGAACTATATGTGCTTCTCTCATTTCTAGTGCCCCVATGRCCATCTACACT 6483
Db 8722 GGCTCTTCAGAACTATATGTGCTTCTCTCATTTCTAGTGCCCCVATGRCCATCTACACT 8781
Qy 6484 GAACAAGATTTATACAGTTATGTGCATATCTAAGCCCGCGCAACAAGAGTACCCATCTT 6543
Db 8782 GAACAAGATTTATACAGTTATGTGCATATCTAAGCCCGCGCAACAAGAGTACCCATCTT 8841
Qy 6544 CCTTTTGTATAGAGCAGGAGTCTAGTGCTAGTACTAGCTAGCTAGCTAGCTAGCTAGCTAG 6603
Db 8842 CCTTTTGTATAGAGCAGGAGTCTAGTGCTAGTACTAGCTAGCTAGCTAGCTAGCTAGCT 8901
Qy 6604 ACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAGCGGTC 6663
Db 8902 ACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAGCGGTC 8961
Qy 6664 GCCGACTCCCTGCTACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGCTCTTCTRA 6723
Db 8962 GCCGACTCCCTGCTACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGCTCTTCTCA 9021
Qy 6724 AATCGAAGAGCTTTAGACTTGTCTAACCCTGAPAGAGGGGGAACCTGTTATTTTAGGG 6783
Db 9022 AATCGAAGAGCTTTAGACTTGTCTAACCCTGGAAGAGGGGGAACCTGTTATTTTAGGG 9081
Qy 6784 GAAGAATGCTGTTATTTATGTTATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db 9082 GAAGAATGCTGTTATTTATGTTATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141
Qy 6844 CSNAGTCGAATCAACGTAAGAGAGCTTTGGAACACTGGAACCTGGGCGCTCTCTC 6903
Db 9142 CGAGATCGAATCAACGTAAGAGAGAGCTTTGGAACACTGGAACCTGGGCGCTCTCTC 9201
Qy 6904 AGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6963
Db 9202 AGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9261
Qy 6964 CTCTCTTTGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7023
Db 9262 CTCTCTTTGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9321
Qy 7024 GCTGTRAAACTCAAAATGAGGCCCAAGATCGAGTCCAGACTAAGATCTACCGCAGACCC 7083
Db 9322 GCTGTRAAACTCAAAATGAGGCCCAAGATCGAGTCCAGACTAAGATCTACCGCAGACCC 9381
Qy 7084 CTGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7143
Db 9382 CTGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9441
Qy 7144 GAAATCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9441
Db 9442 GAAATCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9501
Qy 7204 SGTGCGCAACTCTCCCAACAGCACTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7263
Db 9502 C-TGCGCAACTCTCCCAACAGCACTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9560
Qy 7264 AGGACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7323
Db 9561 AGGACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9620
Qy 7324 ACATCCACTTTAAACAGGGGCTTGAACCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7383
Db 9621 ACATCCACTTTAAACAGGGGCTTGAACCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9680
Qy 7384 ACTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATTTATTGCTG 7443
Db 9681 ACTAAATGCTAATTTAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATTTATTGCTG 9740
Qy 7444 AGAGCAGCAGGAGGGAATGATTCGGGATATAAACCCAAAGTTCAGGCGGCAACGG 7503
Db 9741 AGAGCAGCAGGAGGGAATGATTCGGGATATAAACCCAAAGTTCAGGCGGCAACGG 9800
```

---

```
Qy 7504 CAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db 9801 CAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 9860
Qy 7564 ATTAAATCTTGGCACTGCR 7582
Db 9861 ATTAAATCTTGGCACTGCA 9879

RESULT 13
AX007980
LOCUS 10499 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent WO9967395.
ACCESSION AX007980
VERSION AX007980.1 GI:9995677
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Perin, J. P., Rieger, F. and Alliel, P. M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 129-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
source
1. .10499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 84.9%; Score 6436; DB 6; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

Qy 606 GATGGGAACGTTCCCGCAAGCAAAAGCGCCCTTAAGACCTATTCTGAGAAATGGGA 665
Db 2953 GATGGGAACGTTCCCGCAAGCAAAAGCGCCCTTAAGACCTATTCTGAGAAATGGGA 3012
Qy 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGCACTTATTTCTTCTGAGTCCGCCCTG 725
Db 3013 CCAATTTGACCTCAGACACTAAGAAAGAAAGCACTTATTTCTTCTGAGTCCGCCCTG 3072
Qy 726 GCACCTCTGAGGAGTATAATTAATTAACCACTCTTACAGTAGACTCTTTTGTAGAA 785
Db 3073 GCACCTCTGAGGAGTATAATTAATTAACCACTCTTACAGTAGACTCTTTTGTAGAA 3132
Qy 786 AAGCAATGGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAATCAC 845
Db 3133 AAGCAATGGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAATCAC 3192
Qy 846 AATTATGTAAGAGTGAATTTATGCTCCTACAGGAAGCTTTACAGTCTTACCTCCTATC 905
Db 3193 AATTATGTAAGAGTGAATTTATGCTCCTACAGGAAGCTTTACAGTCTTACCTCCTATC 3252
Qy 906 CCAGCATCCCGACTCTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 965
Db 3253 CCAGCATCCCGACTCTCTCTCCCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
Qy 966 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCATATTATTCCTCAATATGA 1025
Db 3313 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCATATTATTCCTCAATATGA 3372
Qy 1026 CCCCCTCCAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1085
Db 3373 CCCCCT-CCAAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3431
Qy 1086 TCTCCAGAGCTTAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1145
```

Db 3432 TCTCCAGACTTAAGACAAATAAAAA CAGACTTAGGTAAATCTCAGATAACCCCTGATGG 3491  
Qy 1146 CTATATTGRTGTTTTACAGGGTTAGGACAAATCTTTGATCTGACATGAGAGATATATA 1205  
Db 3492 CTAATTATGATCTTTTACAAGGTTAGGACAAATCTTTGATCTGACATGAGAGATATA-A 3550  
Qy 1206 TGTCACTGCTAAATCAGACACTAACCCCAATAGAGAGAGAGTGCACCAATTAATCTGAGCCT 1265  
Db 3551 TGTCACTGCTAAATCAGACACTAACCCCAATAGAGAGAGAGTGCACCAATTAATCTGAGCCT 3610  
Qy 1266 GAGGTGTTGGCGATCTCTGATCTCTCAGTCAAGTCAATGGATANGGATGACACACAGAGG 1325  
Db 3611 GAGAGTTTGGCGATCTCTGATCTCTCAGTCAAGTCAATGGATANGGATGACACACAGAGG 3667  
Qy 1326 AAAGANAATGATTCCTCCACAGCCAGCARGCAGTTCCTCAGTCTASACCTCATTTGGGGAC 1385  
Db 3668 AAAGAGATGATTCCTCCACAGCCAGCARGCAGTTCCTCAGTCTAGACCTCATTTGGG--- 3724  
Qy 1386 ACAGAAATCAGTAACATGGGAGATTTGGTCTGACAGCAATTTGCTAACTTTGTGCTASAA 1445  
Db 3725 ACACAGAAATCAGAAACAT-GGAGATTTGGTCTGCTGACAGCAATTTGCTAACTTTGTGCTAGAA 3783  
Qy 1446 GGACTAAGGAAAACCTASAGAAATCTAYGAATTTACTCAATGATGTCCACCATAACACA 1505  
Db 3784 GGACTAAGGAAAACCTAGGAG-AGTCTATGAATTTACTCAATGATGTCCACCATAACACA 3842  
Qy 1506 GGGGAAGGAAAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGCACTTGAGGAAGC 1565  
Db 3843 -GGGAAGGAAAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGCACTTGAGGAAGC 3901  
Qy 1566 GTGCTCTCTGTCACTGACTCTTCTGAAGGCAACTAATCTTAAGGCTAAGTTTATCA 1625  
Db 3902 GTGCTCTCTGTCACTGACTCTTCTGAAGGCAACTAATCTTAAGGCTAAGTTTATCA 3961  
Qy 1626 CTCAGTCACTGCGACAGACTAG-AAAAAATCTCAAACTCTGCCGTGAGGCGCGAGGCAAA 1684  
Db 3962 CTCAGTCACTGCGACAGACTAGAAAAAATCTCAAACTCTGCCGTGAGGCGCGAGGCAAA 4021  
Qy 1685 ACTTAGAAACCTATTTGAACCTTGGCAACTCGGTTTTTATAATAGAGATCAGAGGAGC 1744  
Db 4022 ACTTAGAAACCTATTTGAACCTTGGCAACTCGGTTTTTATAATAGAGATCAGAGGAGC 4081  
Qy 1745 AGGCGGAACAGGACAAACGGGATTAATAAAGGCCACCGCTTTAGTCAATGACCCCTCAGG 1804  
Db 4082 AGGCGGAACAGGACAAACGGGATTAATAAAGGCCACCGCTTTAGTCAATGACCCCTCAGG 4141  
Qy 1805 CAATGGACTTTGGAGCTCTGGAAAGGGAAGAGCTGGCAATTTGAATGCCCTAATAGG 1864  
Db 4142 CAATGGACTTTGGAGCTCTGGAAAGGGAAGAGCTGGCAATTTGAATGCCCTAATAGG 4201  
Qy 1865 GCTTGTCTCCAGTGGGCTCTACAGGACACTTTAAAAAGATTTGCCAAGTAGAAGTAAG 1924  
Db 4202 GCTTGTCTCCAGTGGGCTCTACAGGACACTTTAAAAAGATTTGCCAAGTAGAAGTAAG 4261  
Qy 1925 CCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG 1984  
Db 4262 CCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG 4321  
Qy 1985 GGACAAAGTCTTTTGTAGTCAAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db 4322 GGACAAAGTCTTTTGTAGTCAAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGG 4381  
Qy 2045 TGCTTGGGGCAAGGCCATCCCATGCCATCACCTCCACAGAGCCCTGGGTATGCTTGACC 2104  
Db 4382 TGCTTGGGGCAAGGCCATCCCATGCCATCACCTCCACAGAGCCCTGGGTATGCTTGACC 4441  
Qy 2105 ATTGAGGCCAGGAGGTTGTCTCTCGGACACTGGTGGCGGCTCTCTTAGTCTTACTCTTC 2164  
Db 4442 ATTGAGGCCAGG-AGGTTGTCTCTCGGACACTGGTGGCGGCTCTCTTAGTCTTACTCTTC 4500  
Qy 2165 TGTCCCGACAACTGTCTCTCGAGATCTGTCACTATTTCTGAGGGGTCCNTTAAGCGGCA 2224  
Db 4501 TGTCCCGACAACTGTCTCTCGAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGCGGGCA 4558

Qy 2225 GTCACTAGATPACTTTTTCAGGCCACTAAGTTATGAATCGGGAGCTTTATTTTTCAC 2284  
Db 4559 GTCACTAGATAC-TTCTCCAGGCCACTAAGTTATG-ACTGGGAGCTTTATTTCTTTTTCAC 4616  
Qy 2285 ATGCTTTTCTAAATATGCTTTGAAAGCCCACTACTCTTTGTTAGGAGAGACATTTCTAGCAA 2344  
Db 4617 ATGCTTTTCTAAATATGCTTTGAAAGCCCACTACTCTTTGTTAGGAGAGACATTTCTAGCAA 4676  
Qy 2345 AAGCAGGGGCGCATTTATACACTCTGAACATAGGAGAGAAACAACCCGTTTGTGTGTTCCCTCG 2404  
Db 4677 AAGCAGGGGCGCATTTATACACTCTGAACATAGGAGAGAGAAACAACCCGTTTGTGT-CCCTCG 4735  
Qy 2405 CTTGAGGAAGAAATTAATCTCTGAGTCTGGGCAACAGAGAGCAATATGACGAGCCAAA 2464  
Db 4736 CTTGAGGAAGAAATTAATCTCTGAGTCTGGGCAACAGAGAGCAATATGACGAG-CAAA 4794  
Qy 2465 GAATGCCGCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAGGCA 2524  
Db 4795 GAATGCCGCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAGGCA 4854  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTCAAAAAGATTCCTTAAGGACTTTAAAGC 2584  
Db 4855 GTACCCCTCAGACCCCAAGGCCCAACAGGACTCCAAAAAGATTCCTTAAGGACTTTAAAGC 4914  
Qy 2585 CCAAGGCTTAGTAAAAACCATGCAATCCCTGCGAGTAATCCGTAAGTGGATGAGGAGG 2644  
Db 4915 CCAAGGCTTAGTAAAAACCATGCAATCCCTGCGAGTACTCC-----AATTTTAGGAG 4967  
Qy 2645 CACAGAAACCCAGTGGACAGTGGAGGTTAGTGAAGATCTCAGGATTTCAATGGAGGC 2704  
Db 4968 TACAGAAACCCCAACAGACAGTGG-GGTTAGTGAAGATCTCAGGATTTATCAAT- GAGGC 5025  
Qy 2705 CGTTGCTCTTTTATACCCAGCTGTACCTAGCCCTTATATCTGTGMYTTCCTCCAAATACCA 2764  
Db 5026 TGTGTTCTCTATAGCGAGCTGTACCTAGCCCTTATCTGTCTTTCCCAATACCA 5085  
Qy 2765 GGAAGCAGAGTGGTTTACSTCTCGGACCTTMAGGATGCCCTTTCTGCACTCCCTGTACA 2824  
Db 5086 GGAAGCAGAGTGGTTTACAGTCTCGGACCTTCAGGATGCCCTTTCTGCACTCCCTGTACA 5145  
Qy 2825 TCCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAACCCARCATCTCACTCAC 2884  
Db 5146 TCCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAACCCCAACATCTCACTCAC 5205  
Qy 2885 CTGCACTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 2944  
Db 5206 CTGCACTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 5265  
Qy 2945 CCAAGACTGAGYCARITYMTCATACCTGGACACTCTTGTCTCTCGTAKGTGGATGATTT 3004  
Db 5266 CCAAGACTGAGCCAACTCTCATACCTGGACA--CTTGTCTCTCGGTAGGTGGATGATTT 5323  
Qy 3005 ACTTTTGGCTGCTVTTTCAGAAACCTTTGCGCATCAAGCCACCCAGGCTCTTMAATTT 3064  
Db 5324 ACTTTTGGCTGCTVTTTCAGAAACCTTTGCGCATCAAGCCCAAGGCTCTTCAATTT 5383  
Qy 3065 CCTCGACCTGCTGCTACAGGTTTCCAAACARARGCTCARCTCTGCTCAGCAGCAGT 3124  
Db 5384 CCTCGACCTGCTGCTACAGGTTTCCAAACARARGCTCARCTCTGCTCAGCAGCAGT 5443  
Qy 3125 TAAATCTTAGGCTTAARATTTATCCAAAGGCAACGAGCCCTCAGTGAAGGAAYATCCA 3184  
Db 5444 ----TACTTAGGCTTAAATTTATCCAAAGGCAACGAGCCCTCAGTGAAGGAACATCCA 5499  
Qy 3185 GCCTATAGTGGCTTATCTCTCATCTCAAAACCTTAAGCACTAAGRBRTTCTTGGGRT 3244  
Db 5500 GCCTATAGTGGCTTATCTCTCATCTCCAAACCTTAAGCACTAAGRBRTTCTTGGGRT 5559  
Qy 3245 AAYAGGTTCTCGCGAAATAGGATTCCTCCAGTGTGGCAATAGCCAGGCTCATTAATA 3304  
Db 5560 AATAGGTTCTCGCGAAATAGGATTCCTCCAGTGTGGCAATAGCCAGGCTCATTAATA 5618

[illegible]

	DB	6683	CGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAACAAGATGGGGAACCTTCACGAGGACAG	6742
	QY	4444	TTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATAACTTTGCCTGCCAATAT	4503
	DB	6743	TTTTCTCCCTCGGACGGTAGCCACTGAAGAGGGAATAACTTTGTGCTGCAACTAT	6802
	QY	4504	CCAATGAAAATTACTTAAAAACCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA	4563
	DB	6803	CCAATGGAANTTACTTAAACCCCTTCATCAAACCTTTCACTTAGGCATCGATAGCACCCA	6862
	QY	4564	TCBARTGGCCAAATCATTTATTTACTGGAACAGGCCCTTTTCAAACCTATCAAGCARATAKT	4623
	DB	6863	TCGATGGCCAAATCATTTATTTACTGGAACAGGCCCTTTCAAACCTATCAAGCAGATAGT	6922
	QY	4624	CAGGGCCTGTGAATGTGCCBARARAATAATCCCCTGCCCTTAATCGCCNAAGCTCTTCAGG	4683
	DB	6923	CAGGGCCTGTGAATGTGCCCAGAGAATAATCCCCTGCCCTTAATCGCCNAAGCTCTTCAGG	6982
	QY	4684	ARAAACAARAACAGGCCATTACCCCTGRAPAARACTGGCAACTGATTTTAAACCAAGCCCC	4743
	DB	6983	AGAACAAAGAACAGGCCATTACCCCTGGAGNAGACTGGCAACTGATTTTAAACCAAGCCCC	7042
	QY	4744	AACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGGTTGGGCARAG	4803
	DB	7043	AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTAGATACCTTTCACGGGTTGGGCAGAG	7102
	QY	4804	GCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAATAA	4863
	DB	7103	GCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAATAA	7162
	QY	4864	ATTCCAGATTTCCGACCTTCCCAGAGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC	4923
	DB	7163	ATTCCAGATTTCCGACCTTCCCAGAGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC	7222
	QY	4924	ACAGTAAACCCAGGGAGTATCCCAAGGGTTAGGTATACGATATCACTTACACTGCGGCTGA	4983
	DB	7223	ACAGTAAACCCAGGGAGTATCCCAAGGGTTAGGTATACGATATCACTTACACTGCGGCTGA	7282
	QY	4984	AGGCCACGTCCTCAGGGAAGTTCGAGAAAATGAATGAAAAYACTCAAAGGACATCTAAAA	5043
	DB	7283	AGGCCACGTCCTCAGGGAAGTTCGAGAAAATGAATGAAAAYACTCAAAGGACATCTAAAA	7342
	QY	5044	AAGCAAAACCCAGAAACCCACCTCACATGSCCTGVTCGTTCGCTATAGCCTTAAANAAGA	5103
	DB	7343	AAGCAAAACCCAGAAACCCACCTCACATGSCCTGVTCGTTCGCTATAGCCTTAAANAAGA	7402
	QY	5104	ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAATATGCTGTATGGAAGGCC	5163
	DB	7403	ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAATATGCTGTATGGAAGGCC	7462
	QY	5164	TTCAATAACCAATGACCTTGCTGTGACCCAGACGCCAACTTGTATGTGAGACATCACT	5223
	DB	7463	TTCAATAACCAATGACCTTGCTGTGACCCAGACGCCAACTTGTATGTGAGACATCACT	7522
	QY	5224	CTTTAGSCAAATATCAACAAGTTCTTAAACAATTAACAGGACCTATCCCTGAGAGAGG	5283
	DB	7523	CTTTAGSCAAATATCAACAAGTTCTTAAACAATTAACAGGACCTATCCCTGAGAGAGG	7582
	QY	5284	GAAAGAACAATTATCCACCCWGTGACATGGTATTAGTCAAAGTCCCTTCVCTCTAATTCOC	5343
	DB	7583	GAAAGAACAATTATCCACCCWGTGACATGGTATTAGTCAAAGTCCCTTCVCTCTAATTCOC	7642
	QY	5344	CATCCCTAGATACATCTCGGGAAGGACCCCTACCCAGTCACTTTATYTATACCCCACTGGG	5403
	DB	7643	CATCCCTAGATACATCTCGGGAAGGACCCCTACCCAGTCACTTTATYTATACCCCACTGGG	7702
	QY	5404	TTAAAGTGCTGAGGTGAGTCTTCGATACATCACTTGAGTCAAAATTCCTGGATCTGTC	5463
	DB	7703	TTAAAGTGCTGAGGTGAGTCTTCGATACATCACTTGAGTCAAAATTCCTGGATCTGTC	7762
	QY	5464	CRAAGGAACCTGAAAATCCAGGAGACAACGCTAGTCTATTCCTGTGAACTCTTAGAGGATT	5523

Db 7763 CAAAGGAACCTGAAATCCAGGAGACAACGCTAGCTATTTCCTGTGAACCTCTTAGAGGAT 7822

Qy 5524 TGGCGCTGCTCTTCAACCAACCAACAGGAGGAAGTAACCTAAATCATAAATCCCCCATG 5583

Db 7823 TGGCGCTGCTCTTCAACCAACCAACAGGAGGAAGTAACCTAAATCATAAATCCCCCATG 7881

Qy 5584 GSCCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTTAYACCTSTTTCACTCTCACTGCA 5643

Db 7882 GCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTTTCACTCTCACTGCA 7941

Qy 5644 CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTCYACOMAGAGTTTCTATGGAGATG 5703

Db 7942 CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCAGAGTTTCTATGGAGATG 8001

Qy 5704 CAGCGTCCGGAATATTGATGCCCCCATCGTATAGAGTCTTSTAGGGGAACCCCAACC 5763

Db 8002 CAGCGTCCGGAATATTGATGCCCCCATCGTATAGAGTCTTCTTAGGGGAACCCCAACC 8061

Qy 5764 TTCACTGCCACACCCATATGSCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 5823

Db 8062 TTCACTGCCACACCCATATGSCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 8121

Qy 5824 GCAAACTCATTTTGGACAGGAATAATGATTAATCTCTAGTTGTCTGAGAGCTTGGGA 5883

Db 8122 GCAAACTCATTTTGGACAGGAATAATGATTAATCTCTAGTTGTCTGAGAGCTTGGGA 8181

Qy 5884 GTCACGTCTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGAGTTCAA 5943

Db 8182 GTCACGTCTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGAGTTCAA 8241

Qy 5944 GATCAGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACCSGGGTACATGCC 6003

Db 8242 GATCAGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACCSGGGTACATGCC 8301

Qy 6004 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCT 6063

Db 8302 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCT 8361

Qy 6064 ACTCGCTGTAGGCTTATTAACCAACCTCTACCTGGGTCCATGAGGTCTCGGCCCAA 6123

Db 8362 ACTCGCTGTAGGCTTATTAACCAACCTCTACCTGGGTCCATGAGGTCTCGGCCCAA 8421

Qy 6124 AACCTACTAACTGTTGGATGCTCCCTCGAACTTCARGCCATATGTTTCAATCCCT 6183

Db 8422 AACCTACTAACTGTTGGATGCTCCCTCGAACTTCARGCCATATGTTTCAATCCCT 8481

Qy 6184 GTACCTGAAACAATGGAAACAATTGAGCAGAGAAATAACCACTTCGCTTTTAGTAGGA 6243

Db 8482 GTACCTGAAACAATGGAAACAATTGAGCAGAGAAATAACCACTTCGCTTTTAGTAGGA 8541

Qy 6244 CCTCTGTTTCCAAATSTGGAATAACCATACCTCAAACTCCTCTGTGTAATAATTTAGC 6303

Db 8542 CCTCTGTTTCCAAATSTGGAATAACCATACCTCAAACTCCTCTGTGTAATAATTTAGC 8601

Qy 6304 AATACTACATACACAACTCCCAATGCACTAGGTGGGTAACTCCTCCACACAAATA 6363

Db 8602 AATACTACATACACAACTCCCAATGCACTAGGTGGGTAACTCCTCCACACAAATA 8661

Qy 6364 GTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTGAT 6423

Db 8662 GTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTGAT 8721

Qy 6424 GGCTCTTCAGATCTATGTGCTTCTCTCATTTCTAGTGCCCCCYATGRCCATCTACACT 6483

Db 8722 GGCTCTTCAGATCTATGTGCTTCTCTCATTTCTAGTGCCCCCYATGRCCATCTACACT 8781

Qy 6484 GAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAACAAAGAGTACCCATCTT 6543

Db 8782 GAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAACAAAGAGTACCCATCTT 8841

Qy 6544 CCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTGCACTAGGTGCGGGTATCAACA 6603

Db 8842 CCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTGCGGGTATCAACA 8901

Qy 6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 6663

Db 8902 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 8961

Qy 6664 GCGGACTCCCTGGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTAGTCTTCRA 6723

Db 8962 GCGGACTCCCTGGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTAGTCTTCRA 9021

Qy 6724 AATCGAAGAGCTTTGACTTGTCTGCTGAGAGAGGGGAACTGTTTATTTTAGGG 6783

Db 9022 AATCGAAGAGCTTTGACTTGTCTGCTGAGAGAGGGGAACTGTTTATTTTAGGG 9081

Qy 6784 GAAGAAATGCTGTTTATTTATTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 6843

Db 9082 GAAGAAATGCTGTTTATTTATTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 9141

Qy 6844 CSNAGTCGAATACAACTGACGAGAGAGCTTCGAAACACCTGGAACCTGGGGCCCTCCTC 6903

Db 9142 CGAGATCGAATACAACTGACGAGAGAGCTTCGAAACACCTGGAACCTGGGGCCCTCCTC 9201

Qy 6904 AGCCBATGATGATGCTGGGATTTCTCCCTCTTAGGACCTCTAGCAGCTATAATATGCTA 6963

Db 9202 AGCCBATGATGATGCTGGGATTTCTCCCTCTTAGGACCTCTAGCAGCTATAATATGCTA 9261

Qy 6964 CTCCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTGTCTCTCCAGAAATCGAA 7023

Db 9262 CTCCTCTTTGGACCTGTATCTTTAACTCTTTTAACTTTGTCTCTCCAGAAATCGAA 9321

Qy 7024 GCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGACTCTACCGCAGACCC 7083

Db 9322 GCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGACTCTACCGCAGACCC 9381

Qy 7084 CTGACCGGCTCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGACCCCTCTGAG 7143

Db 9382 CTGACCGGCTCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGACCCCTCTGAG 9441

Qy 7144 GAAATCTCAGCTGCAACAACCTCTACACGCCCAATTCAGAGAAAGAGTAGAGCGGT 7203

Db 9442 GAAATCTCAGCTGCAACAACCTCTACACGCCCAATTCAGAGAAAGAGTAGAGCGGT 9501

Qy 7204 SGTGGGCAACCTCCCAACAGCAGCAGTCTAGGTTTCTGTTGAGTGGGGACCTGAGAGC 7263

Db 9502 C-TGGGCAACCTCCCAACAGCAGCAGTCTAGGTTTCTGTTGAGTGGGGACCTGAGAGC 9560

Qy 7264 AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGACTCTAGTGGGAGGTGACC 7323

Db 9561 AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGACTCTAGTGGGAGGTGACC 9620

Qy 7324 ACATCCACCTTTAAACAACCGGGCTTGCACTTAGYTCAACCTGACCAATCAGAGAGCTC 7383

Db 9621 ACATCCACCTTTAAACAACCGGGCTTGCACTTAGYTCAACCTGACCAATCAGAGAGCTC 9680

Qy 7384 ACTAAATGCTTAATAGGCAAGAGAGGTAAAGAAATAGCCAAATCATATTTGCGG 7443

Db 9681 ACTAAATGCTTAATAGGCAAGAGAGGTAAAGAAATAGCCAAATCATATTTGCGG 9740

Qy 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTCGAGCCGCAACGG 7503

Db 9741 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTCGAGCCGCAACGG 9800

Qy 7504 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 7563

Db 9801 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT 9860

Qy 7564 ATTAATCTTGCACCTGCR 7582

Db 9861 ATTAATCTTGCACCTGCA 9879

RESULT 14  
AY101586  
LOCUS

AY101586 10229 bp DNA linear PRI 11-FEB-2004

DEFINITION	Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1 locus, allele A, complete sequence.	
ACCESSION	AY101586	
VERSION	AY101586.1 GI:37544407	
KEYWORDS		
SOURCE	Pan troglodytes (chimpanzee)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.	
REFERENCE	1 (bases 1 to 10229)	
AUTHORS	Bonnaud, B., Lucotte, G., Duret, L., and Mandrand, B.	
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
PUBMED	14757826	
REFERENCE	2 (bases 1 to 10229)	
AUTHORS	Mallet, F., Bouton, O. and Oriol, G.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	
FEATURES	Location/Qualifiers	
source	1..10229	
	/organism="Pan troglodytes"	
	/mol_type="genomic DNA"	
	/isolate="1"	
	/isolation_source="PBMC of chimpanzee 1"	
	/db_xref="taxon:9598"	
	/cell_type="PBMC"	
	/notes="isolated by PCR; ERVW1 locus, allele A endogenous_virus: ERV-W"	
LTR	1..779	
mrna	join(248..810,2889..9820)	
	/note="putative mRNA transcript 1; gag-pol pseudogene"	
mrna	join(248..810,7592..9820)	
	/product="envelope glycoprotein"	
	/note="putative mRNA transcript 2"	
mrna	join(248..810,9255..9820)	
	/note="putative mRNA transcript 3; no coding region"	
CDS	7819..9435	
	/note="syncytin"	
	/codon_start=1	
	/product="envelope glycoprotein"	
	/protein_id="AAM68165.1"	
	/db_xref="GI:37544408"	
	translation="MALPVHIFLFTVLPSTLTAPPCCRMSTSSPYQFLRMQRP GNIDAPSVRSKSGTPTTAHTMFRNCHSATLCHNHTHRTWTKIINPSCPGGLV TVCTWYFTHGMSDGGVQDQREKHVEISQLTRVHSRPRYKGLDLKSLHETLR HTRLVSLFTLTGLHEVSAQNPTNCWICLPNRPYVSIPVPEOWNFSTEINTTSV LVGPLVSLNLTHTSNLTKVPSNTTDDLYNYSKPRKRPVILPFVIGAGVLGALG YRCNGSSSCFSLFVPPMTIYEDQLYNYVSKPRKRPVILPFVIGAGVLGALG TGIGITTSDFYKLSOLNGLMDERVADSLVTLQDQLNSLAAYVLQNRALDLITAE RGVCLFLGECCTCYVNGSGIVTEKVKIRRIQRABELRNTGFWGLLSQMWILP PLGLAAILLILFPGCIFNLLVNFVSRISRIEAVKLQMEPKWQSKTKIYRRLDRPAS RSDVNDIKGTPPEILTAQPLLRNSAGSS"	
LTR	9494..10229	
ORIGIN		
Query Match	83.7%; Score 6346.4; DB 9; Length 10229;	
Best Local Similarity	94.9%; Pred. No. 0;	
Matches	6625; Conservative 194; Mismatches 110; Indels 52; Gaps 28;	
QY	606 GATGGAAACGTTCCTCCCAAGACAAACACCCCTTAAGACGTATCTCGGAATTTGGGA	665
DB	2888 GATGGAAACGTTCCTCCCAAGACAAACACCCCTTAAGACGTATCTCGGAATTTGGGA	2947
QY	666 MCAATTTGACCTCAGACACAAAGAAACGACTTATATCTTCAGTGCCTG	725
DB	2948 CCAATTTGACCTCAGACACAAAGAAACGACTTATATCTTCAGTGCCTG	3007
QY	726 GCATCTCTGAGGAGTATAAATTATACACATCTTACAGCTAGACTCTTTTGTAGAA	785
3008	GCATCTCTGAGGAGTATAAATTATACACATCTTACAGCTAGACTCTTTTGTAGAA	3067
QY	786 AAGGCAAAATGAGTGAAGTCCCAATAAGTACAAAATCTTTTTCATTAAGACAACTCAC	845
DB	3068 AAGGCAAAATGAGTGAAGTCCCAATAAGTACAAAATCTTTTTCATTAAGACAACTCGC	3127
QY	846 AATTATGTAATAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	905
DB	3128 AATTATGTAATAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	3187
QY	906 CCAGCATCCCC--GACTCTCTCCCACTCCCACTTAAGGACCCCTTCAACCCCAATGTGCC	963
DB	3188 CCAGCATCCCCCTGACTCTCTCCCACTTAAGGACCCCTTCAACCCCAATGTGCC	3247
QY	964 AAAAGGAGATAGACAAAGGGTAAACAGTGAACAAAGAGTGCCAATATTTCCCAATAT	1023
DB	3248 AAAAGGAGATAGACAAAGGGTAAACAGTGAACAAAGTGTGCCAATATTTCCCGATTAT	3307
QY	1024 GACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCAAGCCAGAGTGCATGTGCTTTT	1083
DB	3308 GACCCCT-CCAAGCAGTGGGAGGAGAAATTCGGCCCAAGCCAGAGTGCATGTGCTTTT	3366
QY	1084 YTTCTCCAGACTTAAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAATCTGAT	1143
DB	3367 TCTCTCTCAGACTTAAAGCAATATAAACAGACCTAGGTAAATTTCTCAGATAACCTGAT	3426
QY	1144 GGTATATATGRTGTTTTTACAAGGGTTAGGACAAATTTTGTGATCTGATGAGAGATATA	1203
DB	3427 GGTATATATGATGTTTTTACAAGGGTTAGGACAAATTTTGTGATCTGATGAGAGATATA	3486
QY	1204 TATGTCACTGCTFAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGAGC	1263
DB	3487 -ATGTCACTGCTFAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGAGC	3545
QY	1264 CTGAGRTTGGCGATCTCTGGTATCTCAGTCAGTCAATGATGATGATGATGATGATGATG	1323
DB	3546 CCAAGAGTTGGCGATCTCTGGTATCTCAGTCAGTCAATG--ATAGGATGACACAG-A	3602
QY	1324 GGAAAGAAATGATTTCCCAACAGCCAGCAGCAGTCTCCAGCTCTASACCTCATTTGGGG	1383
DB	3603 GGAAAGAAATGATTTCCCAACAGCCAGCAGCAGTCTCCAGCTCTAGACCTCATTTGGG-	3661
QY	1384 ACACAGAAATCAGTAAATCGAGAGTTCGTCTGTCAGAGAGTTCATGATGATGATGATG	1443
DB	3662 --ACACAGAAATCAGAGAGTTCGTCTGTCAGAGAGTTCATGATGATGATGATGATG	3718
QY	1444 AAGACTAAGGAAACCTAGAGAGAAATCTATGATGATGATGATGATGATGATGATGATG	1503
DB	3719 AAGACTAAGGAAACCTAGAGAG--AAGCCTATGATGATGATGATGATGATGATGATG	3777
QY	1504 CAGGGGAGGAGAGAAATCTCTGCTCTCTGAGAGAGCTAAGGGAGGAGGATTCAGGAA	1563
DB	3778 CA-GGGAGGAGAGAGAAATCTCTGCTCTCTGAGAGAGCTAAGGGAGGAGGATTCAGGAA	3836
QY	1564 CGGTGCTCTCTGTCACCTGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1623
DB	3837 CGGTGCTCTCTGTCACCTGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3896
QY	1624 CACTCAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1682
DB	3897 CACTCAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3956
QY	1683 AAACCTAGAAACCTTATGAACTTTGGCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG	1742
DB	3957 AAACCTAGAAACCTTATGAACTTTGGCAATTTGAGGAGGAGGAGGAGGAGGAGGAGG	4016
QY	1743 GCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1802
DB	4017 GCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4076
QY	1803 GGCAAGTGGAGCTTTGGAGGCTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1862
DB	4077 GGCAAGTGGAGCTTTGGAGGCTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4136

QY	1863	GGGCTTGCCTCCAGTGGGCTTCAAGAGGACATTTTAAAAAGATTGTCCAAGTAGAAGTA	1922	QY	2943	GCCCAAGACTTGAGYCARITYMTCAFACTTGGAACA	CTTTGTCTCTCRGTAKGTGGATGAT	3002
Db	4137	GGGCTTGCCTCCAGTGGGCTTCAAGAGGACATTTGAAAAAGATTGTCCAAGTAGAAGTA	4196	Db	5203	GCCCAAGACTTGAGGCAATCTCTACACTGGGACATCTTGTCTCGGTAGGTGGATGAT	5262	Db
QY	1923	AGCGGCCCCCTGTCATGCCCCCTTATTTCAAGGGAATCATCTGGAAGGCCCATCTGCCCA	1982	QY	3003	TTACTTTTGGCYGCCYRVTTCAGAAACCTTGTCCCATCAAGCCACCCCAAGCRCTCTTMAAT	3062	QY
Db	4197	AGCGGCCCCCTGTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCATCTGCCCA	4256	Db	5263	TTACTTTTGGCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCCAAGCGCTCTTCAAT	5322	Db
QY	1983	GGGGACAAGAGGCTTTTGTAGTCAGAGCCACTAACACAGATGATCCAGAGCAGGACTGAG	2042	QY	3063	TTCTCGCYACCTGTGGCTTACAWGGTTTCCAAAC SARARGCTCARCTTGTCTCACAGCAG	3122	QY
Db	4257	GGGGACAAGAGGCTTTTGTAGTCAGAGCCACTAACACAGATGATCCAGAGCAGGACTGAG	4316	Db	5323	TTCTCGCTACCTGTGGCTTACATGGTTTCCAAACCAAGGCTCAACTCTGCTCACGGCAG	5382	Db
QY	2043	GGTGCCCTGGGGCAAGGCCCATCCATGCCATCAACCTCTCACAGAGCCCTGGGTATGCTGA	2102	QY	3123	GTTAAATATCTTAGGRTTAARATTTATCCAAAGGCAACARGCCCTCAGTGAAGGAAYRYATC	3182	QY
Db	4317	GGTGCCCTGGGGCAAGGCCCATCCATGCCATCAACCTCTCACAGAGCCCTGGGTATGCTGA	4376	Db	5383	GT-----TACTTAGGGCTAAATTTATCCAAAGGCAACCGAGGACCTCAGTAGGAACAATC	5438	Db
QY	2103	CAATTGAGGCCCAAGGATGTTCTCTCTGAGCACTGGTGGGGTCTTCTTAGTCTTACTCT	2162	QY	3183	CAGCCCTATCTGGCTTATCTCTCATCCAAACCTTAAAGCAATAGGGGATCTCTTGGC	3242	QY
Db	4377	CAATTGAGGCCCAAGG-AGGTTGTCTCTGAGCACTGGTGGGGTCTTCTTAGTCTTACTCT	4435	Db	5439	CAGCCCTATCTGGCTTATCTCTCATCCAAACCTTAAAGCAATAGGGGATCTCTTGGC	5498	Db
QY	2163	TCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGG	2222	QY	3243	RTAAVAGGYTTCTGCCGAATATGGATTCCCGAGGTTGCCRAAATAAGCCAGGYCATTAWA	3302	QY
Db	4436	TCTGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGG	4493	Db	5499	GTAAATAGGTTTCTGCCGATATGGATT-CCAGGTATGGGAAATAGCCAGGTCAITAA	5557	Db
QY	2223	CAGTCACTAGATCTTTTCCAGCCACTAAGTTATGAACTGGGGAGCTTTATCTTTTC	2282	QY	3303	TACASTAATTAAGGAACTCAGAAAGCCAAATCCCATTTTARTTAAGATGGAYAMCTGAAGY	3362	QY
Db	4494	CAGTCACTAGATCT- TTCTCCAGCCACTAAGTTATG- ACTGGGGAGCTTTATCTTTTC	4551	Db	5558	TACACTAATTAAGGAACTCAGAAAGCCAAATCCCATTTTARTTAAGATGGACACTGAAAT	5617	Db
QY	2283	ACATGCTTTTCTAATTATGTTGAAAGCCCCACTAGCTTTGTTAGGGAGACATTTAGC	2342	QY	3363	MRAGTGGCTTTCCAGGCCCTTAAAGAGCCCTTAAACCAAGYCCAGGTGTTAAGVTTG	3422	QY
Db	4552	ACATGCTTTTCTAATTATGCTGAAAGCCCCACTAGCTTTGTTAGGGAGACATTTAGC	4611	Db	5618	AGAGTGGCTTTCCAGGCC-----TAAACAAGCCCCAGGTGTTAAGVTTG	5663	QY
QY	2343	AAAAGCGGGCCATTATACACTGTAGATAGAGAGGACACCCGTTTGTGTGCC	2402	QY	3423	CCAAAGGGCAAGACTTTTSTTTTATYRTCAAGAAAAAACAAGAAAYAGCTCTRGAGT	3482	QY
Db	4612	AAAAGCGGGCCATTATACACTGTAGATAGAGAGGACACCCGTTTGTGTGCC	4670	Db	5664	CCAAAGGGCAAGACTTTTCTCATATGTCAAG-AAAACAAGGAATAGCTTAGGAGT	5722	Db
QY	2403	TGCTTGAAGGAAGTAATTAATCTTGAAGTCTGGGCAACAGAGGACATATGAGCAGCCA	2462	QY	3483	CCTTACACAGRTCCRAGGAYGAGCTTGCAACCYRTGGGRYACCTGASTAAGGAAATGA	3542	QY
Db	4671	TGCTTGAAGGAAGTAATTAATCTTGAAGTCTGGGCAACAGAGGACATATGAGCAG-CA	4729	Db	5723	CCTTACACAGATCCGAGGATGAGCTTGCAACCTGTGGCATACCTGACTAAGGAAATTA	5782	Db
QY	2463	AGAATGCCGCTCTGTTCAAGTTAACTAAAGATTTCACTTCTTCTTCCCTACCAAGG	2522	QY	3543	TGTAGTGGCAAGGGTTGRCYTCATTTTAVGGGTAGTGGTGGAGTAGCAGTYKTAGT	3602	QY
Db	4730	AGAATGCCCATCTGTTTCAAGTTAACTAAAGATTTCACTTCTTCTTCCCTACCAAGG	4789	Db	5783	TGTAGTGGCAAGGGTTGACCTCATTTTATGGGTAGTGGTGGAGTAGCAGCTTTAGT	5842	QY
QY	2523	CAGTACCCCTCAGACCCAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA	2582	QY	3603	ATCTGAAGCAGTTTAAATAATAACAGGAGAGATCTTACTGTGTGGACATCTCATGAGT	3662	QY
Db	4790	CAGTACCCCTCAGACCCAGGCCCAACAGGACTCCAAAGATTTGTTAAGGACTTAA	4849	Db	5843	ATCTGAAGCAGTTTAAATAATAACAGGAGAGATCTTACTGTGTGGACATCTCATGATG	5902	QY
QY	2583	GCCCAAGGCTTAGTAAACCATGATTAATCTCCCTGAGTAACTCCGTAGTGGATTGAGGA	2642	QY	3663	GAAYRGATACCTACCTGCTTAAAGGACCTTGTGGCTGTGAGACAACYGTGTTACTTAATR	3722	QY
Db	4850	GCCCAAGGCTTAGTAAACCATGATTAATCTCCCTGAGTAACTCCGTAGTGGATTGAGGA	4904	Db	5903	GAATGGCATACTCACTGTTTAAAGGAGACTTGTGGCTGTGAGACAACCTGTTTACTTAAATG	5962	QY
QY	2643	GGCAGAGAAACCCAGTGAGCAGTGAGGGGTTAGTGAAGATTTCTCAGGATTTCAATGGAG	2702	QY	3723	TCAGGCTCTATTAATTTGAAGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACC	3782	QY
Db	4905	ATTACAGAAACCCCAACAGACAGTGG- GGTAGTGAAGATCTCAGGATTTCAAT- GAG	4962	QY	5963	TCAGGCTCTATTAATTTGAAGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACC	6022	QY
QY	2703	GCGTGTCTCTTTATACCCAGCTGTACCTAGCCCTTATCTGMYTTTCCCAATACCA	2762	QY	3783	AGYCNCATTTCTTCCAGACAATCAAGAAAAAGATARAAYATAACTGTCAACAARTTATTC	3842	QY
Db	4963	GCTGTGTTCTCTATAGCCAGCTGTACCTAGCCCTTATCTGCTTTTCCCAATACCA	5022	QY	6023	AGCCACATTTCTTCCAGACAATCAAGAAAAAGATAGAACAATACTGTCAACAAGTAAATTC	6082	QY
QY	2763	GAGGAACGAGTGGTTTACASTCCTCGACCTTMAAGATGCTTCTTCTGATCCCTGTA	2822	QY	3843	TCAAAACCTATGCACTCGAGGGGACCTTGTAGAGTTTCTTGTGACTGATCCYGACCTTCA	3902	QY
Db	5023	GAGGAACGAGTGGTTTACAGTCTTGACCTTGACCTTCAGATGGCTTCTTCTGATCCCTGTA	5082	QY	6083	TCAAAACCTATGCACTCGAGGGGACCTTTTGAAGGTTTCTTGTGACTGATCCCGACC- TCA	6141	QY
QY	2823	CATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATCTTCAAAACCCARCATCTCAATC	2882	QY	3903	ACTTGTATCTAGTGGAGTTCTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAG	3962	QY
Db	5083	CATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATCTTCAAAACCCARCATCTCAATC	5142	QY	6142	ACTTGTATCTAGTGGAGTTCTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATGAG	6201	QY
QY	2883	ACCTGGACTTTTACCCCAAGGGTTGAGGATAGYCCCATCTATTTGGCCAGGCAATTA	2942	QY	3963	TGCTCAGTGAATAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTYAGC	4022	QY
Db	5143	ACCTGGACTTTTACCCCAAGGGTTGAGGATAGTCCCCATCTATTTGGCCAGGCAATTA	5202	QY	6202	TGCTCAGGATATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTCAGC	6261	QY
QY				QY	4023	TRCAGAACTAATAGCCYTCAYTKGGGCACTAGATTTAGGAGAGAAAAAGGGYAAATA	4082	QY







HTLVLNFTLTLGLHEVSAQNPTNCWICLPLNFRPVPYIPVPEQWNFSINTTSV  
LVGLVSNLEIHTSNLTCKVFNNTDTTNSQIRWTPPTQIVCLPSGIFVCGTSA  
YCLNGSSSEMSFLSFLVPPMTIYEDLYNVVYVSPKRPVPILPVITGAGVGLG  
TGIGGTTSTOFYKLSOELNGDMERVADSLVTLQDLNSLAARVLOHPRALDILLTAE  
RGCTCLFLGEBECYVYNOISGIVTEKVEIRDRJQRABELRNTGPMGLLSQWMPWILP  
FLGPLAAIILLILFLGFCIFNLLVNFVSSRIEAVLQWEPMPQSKIKIYRPLDRPASP  
RSDVNDIKGTPPBEIILTAOPLLRPNRNSAGS"  
9494..10229

LTR  
ORIGIN

Query Match 83.7%; Score 6344.8; DB 9; Length 10229;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 6624; Conservative 194; Mismatches 111; Indels 52; Gaps 28;

QY	606	GATGGAAAGCTTCCCGCAAGACAAAAAGCCCTTAAGACGATTTCTGGARAAATGGGA	665
DB	2888	GATGGAAAGCTTCCCGCAAGACAAAAAGCCCTTAAGACGATTTCTGGARAAATGGGA	2947
QY	666	MCATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCAGCTG	725
DB	2948	CCAAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCAGCTG	3007
QY	726	GCACCTCAGGGAAGTATAATATAACACATCTTACAGTAGACATCTTTTGTAGAA	785
DB	3008	GCACCTCAGGGAAGTATAATATAACACATCTTACAGTAGACATCTTTTGTAGAA	3067
QY	786	AAGGCAAAATGGAGTGAAGTGCCTAAGTACAAATTTCTTTTCAATTAAGACAACTCAC	845
DB	3068	AAGGCAAAATGGAGTGAAGTGCCTAAGTACAAATTTCTTTTCAATTAAGACAACTCAC	3127
QY	846	AATTAATGTAAGTGTGATTTATGCCCCTACAGGAAGCCCTTACAGTCTACCTCCCTATC	905
DB	3128	AATTAATGTAAGTGTGATTTATGCCCCTACAGGAAGCCCTTACAGTCTACCTCCCTATC	3187
QY	906	CCAGCATCCCC--GACTCTCTCCCAATTAAGACCCCTTCAACCCAAATGGTCC	963
DB	3188	CCAGCATCCCCGACTCTCTCCCAATTAAGACCCCTTCAACCCAAATGGTCC	3247
QY	964	AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATCCCCCAATAT	1023
DB	3248	AAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATCCCCCAATAT	3307
QY	1024	GAACCTTCCAGAGTGGAGGAGAGAAATTCGGCCAGCCAGAGTGCAATGTCYTTTT	1083
DB	3308	GAACCTTCCAGAGTGGAGGAGAGAAATTCGGCCAGCCAGAGTGCAATGTCYTTTT	3366
QY	1084	YVTCCTCCAGACTTAAAGCAATATAACAGACTTAGTAAATTCAGATAATCTGAT	1143
DB	3367	TCCTCTCAGACTTAAAGCAATATAACAGACTTAGTAAATTCAGATAATCTGAT	3426
QY	1144	GGCTATATTGRTGTTTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGATATA	1203
DB	3427	GGCTATATTGRTGTTTTTACAGGGTTAGGACAAATCTTTGATCTGACATGGAGATATA	3486
QY	1204	TATGTCATGCTTAATTCAGACATAACCCAAATGAGAGAAGTGCCACCAATATGTCAGC	1263
DB	3487	-ATGTCATGCTTAATTCAGACATAACCCAAATGAGAGAAGTGCCACCAATATGTCAGC	3545
QY	1264	CTGAGRTTGGCGATCTCTGTATCTCAGTCAATGATGANGATGACAAAGAA	1323
DB	3546	CAAGAGTTGGCGATCTCTGTATCTCAGTCAATGATGANGATGACAAAGAA	3602
QY	1324	GGAAAGANAATGATTTCCCAAGGCGCAGCAGTCTCCAGTCTASACCTCATTTGGG	1383
DB	3603	GGAAAGANAATGATTTCCCAAGGCGCAGCAGTCTCCAGTCTAGACCTCATTTGGG	3661
QY	1384	ACACAGAAATCAGTAACATGGGAGATTTGGTCTGCAGACATTTGCTAACTTGTGTGCTAS	1443
DB	3662	--ACACAGAAATCAGTAACAT--GGAGATTTGGTCTGCAGACATTTGCTAACTTGTGTGCTAG	3718
QY	1444	AAGGCTAAGGAAACTASAGAAARPTCTAATGATTTCTCAATGATGTCACCAATACA	1503
DB	3719	AAGGCTAAGGAAACTAGGAAG--AAGGCTATGAATTAATCTCAATGATGTCACCAATACA	3777

QY	1504	CAGGGAAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGCATTGAGGAA	1563
DB	3778	CA--GGGAAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGCATTGAGGAA	3836
QY	1564	GGGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAACTTCTTAAAGCGTAAGTTAT	1623
DB	3837	GGGTGCTCTCTGTCACTGACTTCTGAGGGCCAACTAACTTCTTAAAGCGTAAGTTAT	3896
QY	1624	CACTAGTCAGCTGAGACATTTAG--AAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAGCA	1682
DB	3897	CACTAGTCAGCTGAGACATTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAGCA	3956
QY	1683	AACTTTAGAAACCCCTATTGAACCTTGCGCACTCTGCTTTTATATATAGAGATCAGAGGA	1742
DB	3957	AACTTTAGAAACCCCTATTGAACCTTGCGTACCTGCTTTTATATATAGAGATCAGAGGA	4016
QY	1743	GCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCCCTTTAGTCATGACCCCTCA	1802
DB	4017	GCAGCGGAAACAGGACAAACAGGATATAAAAAAAGGCCACCCCTTTAGTCATGACCCCTCA	4076
QY	1803	GGCAAGTGAATTTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTCGAATGCCCTAATA	1862
DB	4077	GGCAAGTGAATTTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTCGAATGCCCTAATA	4136
QY	1863	GGGCTTGTCTCAGTGGGCTCTACAGGACACTTTAAAAAAGATTTGTCAGTAGAAGTA	1922
DB	4137	GGGCTTGTCTCAGTGGGCTCTACAGGACACTTTAAAAAAGATTTGTCAGTAGAAGTA	4196
QY	1923	AGCGCCCTCTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA	1982
DB	4197	AGCGCCCTCTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCA	4256
QY	1983	GGGGAACAAAGGTCTTTTGGATCAGAACCACTAACACAGATGATCCAGCAGCAGGACTGAG	2042
DB	4257	GGGGAACAAAGGTCTCTGAGTCAGAACCACTAACACAGATGATCCAGCAGCAGGACTGAG	4316
QY	2043	GGTGTCTGGGCAAGCGCCATCCCATGCTCATCACCCTCACAGAGCCCTGGGTATGCTTGA	2102
DB	4317	GGTGTCTGGGCAAGCGCCATCCCATGCTCATCACCCTCACAGAGCCCTGGGTATGCTTGA	4376
QY	2103	CCATTGAGGGCCAGGAAGGTTGTCTCTGAGACACTGGTGGGCTCTTTAGTCTTACTCT	2162
DB	4377	CCATTGAGGGCCAGG--AGTTTGTCTCTGAGACACTGGTGGGCTCTTTAGTCTTACTCT	4435
QY	2163	TCTGTCCCGGACAACTGCTCTCCAGATCTGTCACTTCTGAGGGGGTCCNTAAGACGGG	2222
DB	4436	TCTGTCCCGGACAACTGCTCTCCAGATCTGTCACTA--TCTGAGGGGGTCC--TAAGACAGG	4493
QY	2223	CAGTCACCTAGATACCTTTTCCAGCCACTAAGTTATGAATGGGGAGCTTTATCTTTTC	2282
DB	4494	CAGTCACCTAGATAC--TTCTCCAGCCACTAAGTTATG--ACTGGGGAGCTTTATCTTTTC	4551
QY	2283	ACATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTTAGGGAGAGACATTTCTAGC	2342
DB	4552	ACATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTTAGGGAGAGACATTTCTAGC	4611
QY	2343	AAAAGCGGGGCCATTTATACCTTGAAACATAGGAAGAAACAACCCGTTTGTGTCNCCC	2402
DB	4612	AAAAGCGGGGCCATTTATACCTTGAAACATAGGAAGAAACAACCCGTTTGTGTCNCCC	4670
QY	2403	TGCTTGGAGGAAGAAATTAATCTCAAGTCTGGGCAACAGAGGACAAATATGACCGAGCA	2462
DB	4671	TGCTTGGAGGAAGAAATTAATCTCAAGTCTGGGCAACAGAGGACAAATATGACCGAG--CA	4729
QY	2463	AAGAAATGCCCTCTGTTTCAAGTTAAATTAAGGANTTCCACTTCTTTCCCTACCAAGG	2522
DB	4730	AAGAAATGCCCTCTGTTTCAAGTTAAATTAAGGANTTCCACTTCTTTCCCTACCAAGG	4789
QY	2523	CAGTACCCCTCTCAGACCCCAAGGCCCAACAGGATTTCAAAAGATTTGTTAAGGACTTAAA	2582
DB	4790	CAGTACCCCTCTCAGACCCCAAGGCCCAACAGGACTTCAAAAGATTTGTTAAGGACTTAAA	4849



Db	6982	CCAAATCTCAGGGATTTTCAGTATCTACTAGTCTGGGTAGATACTTTACGGGTGGGCAG	7041
Qy	4802	AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGTAATAAGSCATCTACTCATGAAA	4861
Db	7042	AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGTAATAAGSCATCTACTCATGAAA	7101
Qy	4862	TAATTCAGAGTTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTCTTTCCAGG	4921
Db	7102	TAATTCAGAGTTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGCCCTCTTTCCAGG	7161
Qy	4922	CCACAGTAACCCAGGAGTATCCACAGGCTTTAGGTATACGATATCACTTACATGCGCCT	4981
Db	7162	CCACAGTAACCCAGGAGTATCCACAGGCTTTAGGTATACGATATCACTTACATGCGCCT	7221
Qy	4982	GAAGGCCACAGTCTCAGGGAAGGTTCGAGAAAATGAATGAAYAACTCAAGAGCATCTAA	5041
Db	7222	GAAGGCCACAGTCTCAGGGAAGGTTCGAGAAAATGAATGAAYAACTCAAGAGCATCTAA	7281
Qy	5042	AAAAAGCAAAACCAGGAACCCACCTCAATGGCCCTGYTCTGTTCCTATAGCCCTTAAAAA	5101
Db	7282	AAAAAGCAAAACCAGGAACCCACCTCAATGGCCCTGYTCTGTTCCTATAGCCCTTAAAAA	7341
Qy	5102	GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	5161
Db	7342	GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	7401
Qy	5162	CCTTCATAACCAATGACTTGTGTGCTTGACCCCAAGCAGCACTTAGTTGCAGACATCAC	5221
Db	7402	CCTTCATAACCAATGACTTGTGTGCTTGACCCCAAGCAGCACTTAGTTGCAGACATCAC	7460
Qy	5222	CTCCTTAGCCAAATPATCAACAAGTTCTTAAACAATTACAAGGAACCTTATCCCTGAGAAGA	5281
Db	7461	CTCCTTAGCCAAATPATCAACAAGTTCTTAAACAATTACAAGGAACCTTATCCCTGAGAAGA	7520
Qy	5282	GGGAAAAGAACTATTCCACCCTTGTGATGATGATGATGATGATGATGATGATGATGATG	5341
Db	7521	GGGAAAAGAACTATTCCACCCTTGTGATGATGATGATGATGATGATGATGATGATGATG	7580
Qy	5342	CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCACTTTATATACCCCACTGC	5401
Db	7581	CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCACTTTATATACCCCACTGC	7640
Qy	5402	GGTTAAAGTGGCTGGAGTCTTGATATACATCACTCGAGTCAAACTCGGATCT	5461
Db	7641	GGTTAAAGTGGCTGGAGTCTTGATATACATCACTCGAGTCAAACTCGGATCT	7700
Qy	5462	GCCAAAGGAACCTGAAATTCAGAGACAACGCTAGCTATTCCTGTGAACTCTTAGAGGA	5521
Db	7701	GCCAAAGGAACCTGAAATTCAGAGACAACGCTAGCTATTCCTGTGAACTCTTAGAGGA	7760
Qy	5522	TTTGGCGCTGCTCTTCAAAACAAACACGAGGGAAGTAAGTAAGTAAGTAAGTAAGTAAG	5581
Db	7761	TTTGGCGCTGCTCTTCAAAACAAACACGAGGGAAGTAAGTAAGTAAGTAAGTAAGTAAG	7819
Qy	5582	TGSGCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCTCTTTTCACTCTCACTG	5641
Db	7820	TGSGCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCTCTTTTCACTCTCACTG	7879
Qy	5642	CACCCCTCCATGCGCTGTATGACAGTACGCTCCCTTACCMAGAGTCTTCTATGAGGAA	5701
Db	7880	CACCCCTCCATGCGCTGTATGACAGTACGCTCCCTTACCMAGAGTCTTCTATGAGGAA	7939
Qy	5702	TGAGAGTCCCGGAAATTTGATGCCCCCATGATAGGATCTTTTAAAGGGAACCCCA	5761
Db	7940	TGAGAGTCCCGGAAATTTGATGCCCCCATGATAGGATCTTTTAAAGGGAACCCCA	7999
Qy	5762	CCTTCATGCCCCACACCATATGCCCCCAACTGCTATCACTCTGCCACTCTTTGATGC	5821
Db	8000	CCTTCATGCCCCACACCATATGCCCCCAACTGCTATCACTCTGCCACTCTTTGATGC	8059
Qy	5822	ATGCAAAATCTCAATTTAGGACAGGAAAATGATTAATCTTAGTGTCTTGGAGGACTTG	5881
Db	8060	ATGCAAAATCTCAATTTAGGACAGGAAAATGATTAATCTTAGTGTCTTGGAGGACTTG	8119
Qy	5882	GAGTCACTGTCTGTGGACTTACTTACCCCAAACTGGTATGTCTGTATGGGGTTCGAGTTC	5941
Db	8120	GAGTCACTGTCTGTGGACTTACTTACCCCAAACTGGTATGTCTGTATGGGGTTCGAGTTC	8179
Qy	5942	AAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG	6001
Db	8180	AAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG	8239
Qy	6002	GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCC	6061
Db	8240	GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCC	8299
Qy	6062	ATACTCGCCCTGTGAAGCTTATTAATACCACTCTCACTGGGTCCATGAGGTCTCGGCC	6121
Db	8300	ATACTCGCCCTGTGAAGCTTATTAATACCACTCTCACTGGGTCCATGAGGTCTCGGCC	8359
Qy	6122	AAAAACCTACTAACTGTGTGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC	6181
Db	8360	AAAAACCTACTAACTGTGTGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC	8419
Qy	6182	CTGTACTGAACTGAAACAACTTCAGCAAGAAATPAAACACACACTTCCGTTTTAGTAG	6241
Db	8420	CTGTACTGAACTGAAACAACTTCAGCAAGAAATPAAACACACACTTCCGTTTTAGTAG	8479
Qy	6242	GACCTCTGTGTTTCCAAATSTGGAATTAACCCATACCTCAAACTCAGCTGTGTAATAATTA	6301
Db	8480	GACCTCTGTGTTTCCAAATSTGGAATTAACCCATACCTCAAACTCAGCTGTGTAATAATTA	8539
Qy	6302	GCATATCTACATACACAACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACACAAA	6361
Db	8540	GCATATCTACATACACAACTCCCAATGCACTCAGGTGGGTAACTCTCCCAACACAAA	8599
Qy	6362	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCTCTATGTTGTTGA	6421
Db	8600	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCTCTATGTTGTTGA	8659
Qy	6422	ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCCCTCCATGRCATCTACA	6481
Db	8660	ATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTAGTGCCCTCCATGRCATCTACA	8719
Qy	6482	CTGAAACAAGATTTATACAGTTATGTCTATCTAAGCCCCCAACAAAGAGTACCCATTC	6541
Db	8720	CTGAAACAAGATTTATACAAATTAATGTCTATCTAAGCCCCCAACAAAGAGTACCCATTC	8779
Qy	6542	TTCTTTTGTATAGAGCAGAGTGTCTAGTGCACTAGGTACTGGCATTTGGCGGTATCA	6601
Db	8780	TTCTTTTGTATAGAGCAGAGTGTCTAGTGCACTAGGTACTGGCATTTGGCGGTATCA	8839
Qy	6602	CAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGGAATGGAACGGG	6661
Db	8840	CAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGGAATGGAACGGG	8899
Qy	6662	TGCGCACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGAGTATGCTTTC	6721
Db	8900	TGCGCACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGAGTATGCTTTC	8959
Qy	6722	RAAATCGAAGAGCTTTAGACTGCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAG	6781
Db	8960	RAAATCGAAGAGCTTTAGACTGCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAG	9019
Qy	6782	GGGAGAAATGCTGTTTATATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	6841
Db	9020	GGGAGAAATGCTGTTTATATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	9079
Qy	6842	TTCSAGATCGAATACAACTAGCAGAGAGAGCTTCGAAACACTGGACCTCGGGCCTCC	6901
Db	9080	TTCSAGATCGAATACAACTAGCAGAGAGAGCTTCGAAACACTGGACCTCGGGCCTCC	9139
Qy	6902	TCAGCCATATGATGCTTCCCTTTCTTAGGAACTCTTAGGAACTCTTAGGAACTCTTAGG	6961
Db	9140	TCAGCCATATGATGCTTCCCTTTCTTAGGAACTCTTAGGAACTCTTAGGAACTCTTAGG	9199

```
Qy 6962 TACTCCTCTTTGGACCCCTGTATCTTTTACCTCCTTGTAACTTTGTCTCTTCCAGAAATCG 7021
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9200 TACTCCTCTTTGGACCCCTGTATCTTTTAACTCCTCTTGTAACTTTGTCTCTTCCAGAAATCG 9259
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7022 AAGCTGTAAACTTAACTAAGAGCCCAAGATGCAAGTCAAGACTAAGATCTACCGCAGAC 7081
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9260 AAGCTGTAAACTTAACTAAGAGCCCAAGATGCAAGTCAAGACTAAGATCTACCGCAGAC 9319
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7082 CCCTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGATCATCAAGGACACCCCTCCTG 7141
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9320 CCCTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGATCATCAAGGACACCCCTCCTG 9379
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7142 AGGAAATCTCAGCTGCACTCACTTACTAGCCCCCAATTCAGCAGGAGCAGTTAGAGCG 7201
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9380 AGGAAATCTTAACTGCACTCACTTACTAGCCCCCAATTCAGCAGGAGCAGTTAGAGCG 9439
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7202 GTSCTCGGCCAACCTCCGCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGAG 7261
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9440 GTCGTCGGCCAACTCCGCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGAG 9499
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7262 ACAGGACTAGCTGATTTCTTAGGCTGATTAAGATCCYTAAGCCTAGSTGGGAAGTGA 7321
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9500 ACAGGACTAGCTGATTTCTTAGGCTGATTAAGATCCYTAAGCCTAGSTGGGAAGTGA 9559
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7322 CCACATCCACCTTTAAACACCGGGCTTGCACTTAGYTCACTGACCAATCAGAGAGC 7381
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9560 CCACATCCACCTTTAAACACCGGGCTTGCACTTAGYTCACTGACCAATCAGAGAGC 9619
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7382 TCACATAAATGCTAATTAGGCAAGACAGAGGTAAGAAATAGCCCAATCATYTTATTGCM 7441
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9620 TCACATAAATGCTAATTAGGCAAGAAACAGAGGTAAGAAATAGCCCAATCATCTATTGCC 9679
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7442 TGAGAGCACAGCAGGAGGCAATGATCGGGATATTAACCAAGTTCGAGCCGCAAC 7501
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9680 TGRAGAGCACAGCAGGAGGCAATGATCGGGATATTAACCAAGTTCGAGCTGGCAAC 9739
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7502 GGCAACCCCTTTGGGTCCCTTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCACCT 7561
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9740 GGCAACCCCTTTGGGTCCCTTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCACCT 9799
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 7562 CTATTAAATCTTGCACTGCR 7582
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 9800 CTATTAAATCTTGCACTGCA 9820
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: November 16, 2005, 16:57:00  
Job time : 31274 secs





[illegible]

5979	GGAAAAAGTCTCTCTGAGTCAAGGCCATAACCAAGATGATCCAGCAGCAGGACTGAGGG	6038
2045	TGCCTGGGGCAAGCGCATCCCATGCGATCACCCCTCACAGAGCCCTGGGTATGCTTGACCC	2104
6039	TGCCTGGGGCAAGCGCATCCCATGCGATCACCCCTCACAGAGCCCTGGGTATGCTTGACCC	6098
2105	ATTGAGGGCCAGGAAGGTTGTCTCTCGGACACTGCTGGACACTGCTGGCGGTCTTCTAGTCTTACTCTTC	2164
6099	ATTGAGGGCCAGG - AGGTTGTCTCTCGGACACTGCTGGCGGTCTTCTAGTCTTACTCTTC	6157
2165	TGTTCCCGGCAACTGCTCTCCAGATCTGTCACTATTCTTGAGGGGGTCCNTAAGACGGGCA	2224
6158	TGTTCCCGGCAACTGCTCTCCAGATCTGTCACTA - TCTGAGGGGGTCC - TAAAGACGGGCA	6215
2225	GTCACTAGATACCTTTTCCAGCGCACTAAAGTTATGAACCTGGGGAGCTTTATCTTTTCAC	2284
6216	GTCACTAGATAC - TTCTCTCCAGCGCACTAAGTTATG - ACTGGGGAGCTTTATCTTTTCAC	6273
2285	ATGCTTTTCTAAATTATGCTTTGAAAGCCCCATACTTGTGTAGGGAGAGACATTTCTAGCAA	2344
6274	ATGCTTTTCTAAATTATGCTTTGAAAGCCCCACTACCTTGTGTAGGGAGAGACATTTCTAGCAA	6333
2345	AAGCAGGGGCCATTATACACCTGCAACATAGGAGAAAGACACCGCTTGTGTGTCGCCCTG	2404
6334	AAGCAGGGGCCATTATACACCTGAACATAGGAGAAAGAACCCCGTTGTGTGT - CCCCTG	6392
2405	CTTGAGGAAGGAATTAATCCTGAAAGCTGGGCAACAGAAAGCAATAATGACGAGCCAAA	2464
6393	CTTGAGGAAGGAATTAATCCTGAAAGCTGGGCAACAGAAAGCAANTATGACGAG - CAAA	6451
2465	GAATGCCGCTCTGTTCAAGTTAAAATAAAGGATTTCACTTCTTTCCCTTACCAAGGCA	2524
6452	GAATGCCGCTCTGTTCAAGTTAAAATAAAGGATTTCCACCTCTCTTCCCTTACCAAGGCA	6511
2525	GTACCCCTCTCAGACCCCAAGGCCCAAGAGATTTCCAAAGATTTGTTAAGGACTTTAAAGC	2584
6512	GTACCCCTCTCAGACCCCAAGGCCCAAGAGACTTCCAAAGATTTGTTAAGGACTTAAAGC	6571
2585	CCAAGCTTAGTAAAAACCATGATACCTCCCTGCAGTAAATTCGGTAGTGATGAGGAGG	2644
6572	CCAAGCTTAGTAAAAACCATGATACCTCCCTGCAGTACTCC - - - - - AATTTTAGGAG	6624
2645	CACAGAAACCCAGTGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTTCAATGAGGCG	2704
6625	TACAGAAACCCACAGACAGTGA - GGTTAGTGCAAGATCTCAGGATTTCAAT - GAGGC	6682
2705	CGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAAAATACCAGA	2764
6683	TGTTGTTCTCTATAGCCAGCTGTACTAGCCCTTATCTGTGCTGCTTCCCAAAATACCAGA	6742
2765	GGAAGCAGAGTGGTTTACAGTCTCTGGACCTTMAAGGATGCCTTCTCTGCAATCCCTGTACA	2824
6743	GGAAGCAGAGTGGTTTACAGTCTCTGGACCTTMAAGGATGCCTTCTCTGCAATCCCTGTACA	6802
2825	TCCTGACTCTCAATCTTGTGTTGCTTTGAAGATPACTTCAAACCCACATCTCAACTCAC	2884
6803	TCCTGACTCTCAATCTTGTGTTGCTTTGAAGATPACTTCAAACCCACATCTCAACTCAC	6862
2885	CTGGACTRTTTTACCCCAAGGGTTCAGGATAGVCCCCATCTATTTTGGCCAGGCATTAGC	2944
6863	CTGGACTRTTTTACCCCAAGGGTTCAGGATAGTCCCCATCTATTTTGGCCAGGCATTAGC	6922
2945	CCAAGACTTCGAGTCARTTMTCATACCTGGACACTCTTGTGCTTCRCGTAKGTGGATGATTT	3004
6923	CCAAGACTTCGAGTCARTTMTCATACCTGGAC - -CTTGTCTTCGGTAGGTGGATGATTT	6988
3005	ACTTTTTCGCTCCCYRTTCAGAAAACCTTTGTGCCATCAAGCCCAACCAAGRCCTCTTMAATTT	3064
6981	ACTTTTTCGCTCCCYRTTCAGAAAACCTTTGTGCCATCAAGCCCAACCAAGRCCTCTTMAATTT	7044
3065	CTTCGCACCTGTGGCTACAGWGTTCCTCAACCSNARGCTCARTCTGCTCTCACAGCAGGT	3124
7041	CTTCGCCTACTGTGGCTACATGGTTTCCAAACCAAAAGGCTCACTCTGCTCCACAGCAGGT	7104



Db	9240	GAAGAACTATTTCCACCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAAATCC	9299	Db	10319	GTCTGCTACCTCAGGAATAATTTTGTCTGTGATACTCAGCCTATCGTTTGAAT	10378
Qy	5344	CATCCCTAGATACATCTCTGGAAGGACCTTACCAGTCATTTATATACCCCAACTGCGG	5403	Qy	6424	GGCTCTTCAAGATCTATGTGCTTCTCTCATTTAGTGCCTCCCTCATTTAGTGCCTCCCTCATCTT	6483
Db	9300	CATCCCTAGATACATCTCTGGAAGGACCTTACCAGTCATTTATATACCCCAACTGCGG	9359	Db	10379	GGCTCTTCAAGATCTATGTGCTTCTCTCATTTAGTGCCTCCCTCATTTAGTGCCTCCCTCATCTT	10438
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGGAGTCAAAATCCTGGATCTGC	5463	Qy	6484	GAACAAGATTTATACAGTTATGTCTATCTAAGACCCCGCAACAAAGAGTACCCATTCCT	6543
Db	9360	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGGAGTCAAAATCCTGGATCTGC	9419	Db	10439	GAACAAGATTTATACAGTTATGTCTATCTAAGACCCCGCAACAAAGAGTACCCATTCCT	10498
Qy	5464	CAAAAGAACTGAAATTCAGAGAGACAAACCTAGCTATTCTCTGTGAACCTCTAGAGGATT	5523	Qy	6544	CCTTTTGTATAGGACAGGAGTCTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCAT	6603
Db	9420	CAAAAGAACTGAAATTCAGAGAGACAAACCTAGCTATTCTCTGTGAACCTCTAGAGGATT	9479	Db	10499	CCTTTTGTATAGGACAGGAGTCTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCAT	10558
Qy	5524	TGCGCTGCTCTTCAACACACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG	5583	Qy	6604	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC	6663
Db	9480	TGCGCTGCTCTTCAACACACACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG	9538	Db	10559	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC	10618
Qy	5584	GSCTCTCCCTTATCATATTTTCTCTTASTGTSTTTTACCTSTTTTCACTCTCACTGCA	5643	Qy	6664	GCGACTCCTCGTCACTCTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCTCTTCA	6723
Db	9539	GSCTCTCCCTTATCATATTTTCTCTTASTGTSTTTTACCTSTTTTCACTCTCACTGCA	9598	Db	10619	GCGACTCCTCGTCACTCTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCTCTTCA	10678
Qy	5644	CCCCCTCCATGCGCTGATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAATG	5703	Qy	6724	AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGGAACTGTTTATTTTAGGG	6783
Db	9599	CCCCCTCCATGCGCTGATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAATG	9658	Db	10679	AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGGAACTGTTTATTTTAGGG	10738
Qy	5704	CAGGCTCCCGAAATATTGATGCCCATCTATAGGATCTTTSTAAGGAAACCCCATC	5763	Qy	6784	GAAGAATGCTGTTTATTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	6843
Db	9659	CAGGCTCCCGAAATATTGATGCCCATCTATAGGATCTTTSTAAGGAAACCCCATC	9718	Db	10739	GAAGAATGCTGTTTATTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	10798
Qy	5764	TTCACTGTCACACACCATATGCCCCGCAACTGCTATCACTCTGCGCATCTTTGCACTGCAT	5823	Qy	6844	CSAGATCGAATCAACCTGATAGCAGAGAGCTTCGAAACAACCTGAGCCCTGGGCTCCTC	6903
Db	9719	TTCACTGTCACACACCATATGCCCCGCAACTGCTATCACTCTGCGCATCTTTGCACTGCAT	9778	Db	10799	CGAGATCGAATCAACCTGATAGCAGAGAGCTTCGAAACAACCTGAGCCCTGGGCTCCTC	10858
Qy	5824	GCAAAATCTCATTTTGGACAGGAAATGATTAATCTAGTGTCTGAGAGACTTGGGA	5883	Qy	6904	AGCCATGAGTGCCTGGATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTCCTA	6963
Db	9779	GCAAAATCTCATTTTGGACAGGAAATGATTAATCTAGTGTCTGAGAGACTTGGGA	9838	Db	10859	AGCCATGAGTGCCTGGATCTCCCTCTTAGGACCTCTAGCAGCTATAATATTCCTA	10918
Qy	5884	GTCACTGCTGTGGACTTCTTACCCCAACTGGTATGTCATGGGGTGGAGTTCAA	5943	Qy	6964	CTCCTCTTGGACCTGTATCTTTTACCTCTTGTAACTTTCTCTCTCCAGAAATCGAA	7023
Db	9839	GTCACTGCTGTGGACTTCTTACCCCAACTGGTATGTCATGGGGTGGAGTTCAA	9898	Db	10919	CTCCTCTTGGACCTGTATCTTTTACCTCTTGTAACTTTCTCTCTCCAGAAATCGAA	10978
Qy	5944	GATCAGCAAGAGAAACATGTAAAGAAAGTAACTCTCCCAACTCAACCCGGGTACATGGC	6003	Qy	7024	GCTGTAAACTCAAAATGAGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC	7083
Db	9899	GATCAGCAAGAGAAACATGTAAAGAAAGTAACTCTCCCAACTCAACCCGGGTACATGGC	9958	Db	10979	GCTGTAAACTCAAAATGAGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC	11038
Qy	6004	ACCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTAACCAT	6063	Qy	7084	CTGGAACCGGCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGACCCCTCCCTGAG	7143
Db	9959	ACCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTAACCAT	10018	Db	11039	CTGGAACCGGCTGTAGCCCAACGATCTGATGTTTAAATGACATCAAGGACCCCTCCCTGAG	11098
Qy	6064	ACTGCGCTGTAAGCTTATTAATACACCTCACTGGGCTCCATGAGGTCTCGGCCAA	6123	Qy	7144	GAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT	7203
Db	10019	ACTGCGCTGTAAGCTTATTAATACACCTCACTGGGCTCCATGAGGTCTCGGCCAA	10078	Db	11099	GAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT	11158
Qy	6124	AACCTACTAATCTGTTGATGCTCCCTCGAATCTTCAAGCCATATGTTTCAATCCCT	6183	Qy	7204	SGTGGCCCAACCTCCCAACAGCAGACTTAGGTTTTCTGTTTGGATGGGGGACCTGAGAGAC	7263
Db	10079	AACCTACTAATCTGTTGATGCTCCCTCGAATCTTCAAGCCATATGTTTCAATCCCT	10138	Db	11159	CGTGGCCCAACCTCCCAACAGCAGACTTAGGTTTTCTGTTTGGATGGGGGACCTGAGAGAC	11218
Qy	6184	GTACCTGAACAATGGAAACAATTCAGACAGAAATAAACCACTTCCGTTTTAGTAGGA	6243	Qy	7264	AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCCTTAAGCCCTAGSTGGAGGTCAC	7323
Db	10139	GTACCTGAACAATGGAAACAATTCAGACAGAAATAAACCACTTCCGTTTTAGTAGGA	10198	Db	11219	AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCCTTAAGCCCTAGSTGGAGGTCAC	11278
Qy	6244	CCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACTCCTGTTGAAATTTAGC	6303	Qy	7324	ACATCCACCTTTTAAACACCGGGCTTGCACTTAGTCTACACCTGACCAATCAGAGAGCTC	7383
Db	10199	CCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACTCCTGTTGAAATTTAGC	10258	Db	11279	ACATCCACCTTTTAAACACCGGGCTTGCACTTAGTCTACACCTGACCAATCAGAGAGCTC	11338
Qy	6304	AATACTACATACACCAACTCCCAATGATCAGGGGGGTAACTCTCCCAACACAAATA	6363	Qy	7384	ACTAAATGCTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTATTCGWTG	7443
Db	10259	AATACTACATACACCAACTCCCAATGATCAGGGGGGTAACTCTCCCAACACAAATA	10318	Db	11339	ACTAAATGCTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTATTCGWTG	11398
Qy	6364	GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTTGAAT	6423	Qy	7444	AGAGCAGCAGGAGGGAACAATGATGGGATATAACCCCAAGTCTTCGAGCCCGCAACGG	7503
				Db	11399	AGAGCAGCAGGAGGGAACAATGATGGGATATAACCCCAAGTCTTCGAGCCCGCAACGG	11458

QY 7504 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563  
 DB 11459 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTCACTCT 11518

QY 7564 ATTAATCTTGCARTGCR 7582  
 DB 11519 ATTAATCTTGCARTGCA 11537

RESULT 2  
 US-09-573-080A-21  
 ; Sequence 21, Application US/09573080A  
 ; Patent No. 6828097  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL  
 ; APPLICANT: ROGAN, PETER  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/09/573,080A  
 ; CURRENT FILING DATE: 2000-05-16  
 ; NUMBER OF SEQ ID NOS: 479  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat\_region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 ; PUBLICATION INFORMATION:  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Jurka, J; Malchiewicz, J; Milosavljevic, A  
 ; TITLE: Prototypic sequences for human repetitive DNA  
 ; JOURNAL: Journal of Molecular Evolution  
 ; VOLUME: 35  
 ; ISSUE: 4  
 ; PAGES: 286-291  
 ; DATE: 1992-10-  
 ; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
 ; DATABASE ENTRY DATE: 1996-01-26  
 ; DATABASE ENTRY DATE: 1996-01-26

Query Match 80.8%; Score 6124.6; DB 4; Length 8523;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGAAACGTTCCCGCAAGACAAACCGCCCTAAGACGTATTCTGGARAATTGGGA 665  
 DB 1923 GATGGAAACGTTCCCGCAAGACAAACCGCCCTAAGACGTATTCTGGARAATTGGGA 1982

QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAAAGCACTATATTCTTGTGAGTCGCGCTG 725  
 DB 1983 CCAATTTGACCCCTCAGACACTAAGAAAGAAAGCACTATATTCTTGTGAGTCGCGCTG 2042

QY 726 GCATCTCTGGGGAAGTAAATATATACACCATCTTACAGCTAGACTCTTTTGTAGAA 785  
 DB 2043 GCATCTCTGGGGAAGTAAATATATACACCATCTTACAGCTAGACTCTTTTGTAGAA 2102

QY 786 AAGCAATGAGTGAAGTGCCATTAAGTACAAACCTTTCTTTTCAATTAAGACAACTCAC 845  
 DB 2103 AAGCAATGAGTGAAGTGCCATTAAGTACAAACCTTTCTTTTCAATTAAGACAACTCAC 2162

QY 846 AATTATGTAAAGTGTGATTATGCTTCCCTCAGGAAGCTTTCAGAGTCTACCTCCCTATC 905  
 DB 2163 AATTATGTAAAGTGTGATTATGCTTCCCTCAGGAAGCTTTCAGAGTCTACCTCCCTATC 2222

QY 906 CCAGCATCCCGACTCTCTTCCCACTTAAAGGACCCCTTCAACCCAAATGGTCAA 965  
 DB 2223 CCAGCATCCCGACTCTCTTCCCACTTAAAGGACCCCTTCAACCCAAATGGTCAA 2282

966 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATATGA 1025  
 DB 2283 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATATGA 2342

QY 1026 CCCCTCCCAAGCAGTGGAGGAGAGAAATTCGGGCCAGCCAGAGTGCATGCTGCTTTT 1085  
 DB 2343 CCCCT-CCCAAGCAGTGGAGGAGAGAAATTCGGGCCAGCCAGAGTGCATGCTGCTTTT 2401

QY 1086 TCTCCAGACTTAAAGCAAAATATAAAGAGAGCTTAAATTTCTCAGATAATCTGATGG 1145  
 DB 2402 TCTCCAGACTTAAAGCAAAATATAAAGAGAGCTTAAATTTCTCAGATAATCTGATGG 2461

QY 1146 CTATATTGRTGTTTACAAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATA 1205  
 DB 2462 CTATATTGATGTTTACAAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGATATA-A 2520

QY 1206 TGTCACTGTCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCACCAATATCTGAGCCT 1265  
 DB 2521 TGTCACTGTCTAAATCAGACACTAAACCCCAATAGAGAGAGTGCACCAATATCTGAGCCT 2580

QY 1266 GAGGTTTGGGATCTCTGTTATCTCAGTCAGTCAATGATANGGATGACCAACAGAGG 1325  
 DB 2581 GAGGTTTGGGATCTCTGTTATCTCAGTCAGTCAATG--ATAGGATGACCAACAG-AGG 2637

QY 1326 AAAGANAATGATTTCCCAAGGCCAGCAGCTTCCAGTCTASACCTCATTTGGGGAC 1385  
 DB 2638 AAAGAGATGATTTCCCAAGGCCAGCAGCTTCCAGTCTAGACCTCATTTGGG--- 2694

QY 1386 ACAGAAATCAGTAACATGGGAGATTTGTTGTCAGACATTTGCTTAACTTTGTTGCTASAA 1445  
 DB 2695 ACAGAAATCAGTAACAT--GGAGATTTGTTGTCAGACATTTGCTTAACTTTGTTGCTAGAA 2753

QY 1446 GGACTAAGGAAACCTAGGAGAAATCTAYCAATTTACTCAATGATGTCACCATTAACACA 1505  
 DB 2754 GGACTAAGGAAACCTAGGAG--AAGTCTATGAAATTTACTCAATGATGTCACCATTAACACA 2812

QY 1506 GGGGAAGGAGAAATCTCTACTGCTTTCTGGAGAGACTTAAGGAGGAGCATTTGAGGAAGC 1565  
 DB 2813 -GGGAAGGAGAAATCTCTACTGCTTTCTGGAGAGACTTAAGGAGGAGCATTTGAGGAAGC 2871

QY 1566 GTGCTCTCTGTCACCTGACTCTTTCTGAAGGCCAACTTAATCTTAAAGCGTAAAGTTTATCA 1625  
 DB 2872 GTGCTCTCTGTCACCTGACTCTTTCTGAAGGCCAACTTAATCTTAAAGCGTAAAGTTTATCA 2931

QY 1626 CTGAGTCAGCTGCAGACATTTAG-AAAAAATCTCAAAAGTCTGCCGTAGGCGCCGAGCAAA 1684  
 DB 2932 CTGAGTCAGCTGCAGACATTTAGAAAAAATCTCAAAAGTCTGCCGTAGGCGCCGAGCAAA 2991

QY 1685 ACTTAGAAACCTTATTGAACTTTGGCAACCTTCGGTTTCTTTTATAATAGAGATCAGGAGGAGC 1744  
 DB 2992 ACTTAGAAACCTTATTGAACTTTGGCAACCTTCGGTTTCTTTTATAATAGAGATCAGGAGGAGC 3051

QY 1745 AGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGG 1804  
 DB 3052 AGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGG 3111

QY 1805 CAAGTGGACTTTGGAGCTCTGGAAGGAGAAAGCTGGGCAATTTGATTCCTTAATAGG 1864  
 DB 3112 CAAGTGGACTTTGGAGCTCTGGAAGGAGAAAGCTGGGCAATTTGATTCCTTAATAGG 3171

QY 1865 GCTTGTCTCAGTCCGCTCTACAGGACACTTTTAAAAAAGATTTGTCAGGTAGAGTAAGTAAG 1924  
 DB 3172 GCTTGTCTCAGTCCGCTCTACAGGACACTTTTAAAAAAGATTTGTCAGGTAGAGTAAGTAAG 3231

QY 1925 CCGCCCTTCTGTCATGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG 1984  
 DB 3232 CCGCCCTTCTGTCATGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGG 3291

QY 1985 GGCACAAAGGCTTTTGTGATCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGG 2044  
 DB 3292 GGCACAAAGGCTTTCTGATCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGG 3351

QY 2045 TGCTGGGGCAAGCGCCATCCCATGCCATCACCTCTCAGAGGCCCTGGGTATGCTTGACC 2104

Db 3352 TGCCTGGGCAAGCGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 3411  
Qy 2105 ATTGAGGGCAGGAAGGTTGTCTCCCTGGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC 2164  
Db 3412 ATTGAGGGCAGG-AGGTTGTCTCTGGACACTGGTGGCGTCTTCTTAGTCTTACTCTTC 3470  
Qy 2165 TGTCCCGGCAACTGCTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNNTAAGACGGCA 2224  
Db 3471 TGTCCCGGCAACTGCTCTCCAGATCTGTCACTA-TCCTGAGGGGGTCC-TAAGACGGCA 3528  
Qy 2225 GTCACATAGATACCTTTTCCAGGCACTAAGATTATGAATCTGGGAGCTTTATTCTTTTCAC 2284  
Db 3529 GTCACATAGATAC-TTCTCCAGGCACTAAGTTATG-ACTGGGGAGCTTTATTCTTTTCAC 3586  
Qy 2285 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGAGAGACATCTAGCAA 2344  
Db 3587 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGAGAGACATCTAGCAA 3646  
Qy 2345 AAGCAGGGCCATTATACACCTGAACATAGGAAGGAACACCCGTTTGTGTCNCCCTG 2404  
Db 3647 AAGCAGGGCCATTATACACCTGAACATAGGAAGGAACACCCGTTTGTGT-CCCTG 3705  
Qy 2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGAACAATATGGACAGCAAA 2464  
Db 3706 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGAACAATATGGACAG-CAAA 3764  
Qy 2465 GAATGCCGCTCTGTTCAAGTTAACTAAGGATTCACCTCTCTTCCCTACCAAGGCA 2524  
Db 3765 GAATGCCGCTCTGTTCAAGTTAACTAAGGATTCACCTCTCTTCCCTACCAAGGCA 3824  
Qy 2525 GTACCCCTCAGACCCCAAGGCCAAACAAGGATTCAAAAGATTGTTAAGGACTTAAAGC 2584  
Db 3825 GTACCCCTCAGACCCCAAGGCCAAACAAGGATTCAAAAGATTGTTAAGGACTTAAAGC 3884  
Qy 2585 CCAAGGCTTAGTAAACCAATGCATAACTCCCTGCACTAATCCCTGATGATTTGAGAGG 2644  
Db 3885 CCAAGGCTTAGTAAACCAATGCATACCTCCCTGCACTAATCCCTGATGATTTGAGAG 3937  
Qy 2645 CACAGAAACCAAGTGGACAGTGGAGGTTAGTCCAGATCTCAGGATTCAATGAGGC 2704  
Db 3938 TACAGAAACCAACAGACAGTGA-GGTTAGTGAAGATCTCAGGATTATCAAT-GAGGC 3995  
Qy 2705 CGTGTCTCTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATATACAGA 2764  
Db 3996 TGTGTCTCTATAGCCAGCTGTACCTAGCCCTTATCTGTCTGCCCAATATACAGA 4055  
Qy 2765 GGAAGCAGAGTGGTTACATCTCTGGACCTTCAAGATGCCCTTCTGTGCATCCCTGTACA 2824  
Db 4056 GGAAGCAGAGTGGTTACATCTCTGGACCTTCAAGATGCCCTTCTGTGCATCCCTGTACA 4115  
Qy 2825 TCCTGACTCTCAATTCTGTTGGCTTTGAAGATCTTCAACCCCAATCTCAACTCAC 2884  
Db 4116 TCCTGACTCTCAATTCTGTTGGCTTTGAAGATCTTCAACCCCAATCTCAACTCAC 4175  
Qy 2885 CTGACTCTTTTACCCCAAGGTTTCAAGGATAGTCCCATCTATTGTCGAGGATTTAGC 2944  
Db 4176 CTGACTTTTACCCCAAGGTTTCAAGGATAGTCCCATCTATTGTCGAGGATTTAGC 4235  
Qy 2945 CCAAGACTCTGAGYCARITYMTACATCTGGACACTCTTGTCTTCRGTAGTGGATTTT 3004  
Db 4236 CCAAGACTCTGAGCCCAATCTCATACCTGGACA--CTTGTCTTCGGTAGTGGATTTT 4293  
Qy 3005 ACTTTTGGYCCVRRTTTCAAGAACCTTGTGGCATCAAGCCACCCCAAGCRCTCTTMAATTT 3064  
Db 4294 ACTTTTGGCGGCCAATTCAGAAACCTTGTGCAATCAAGCCACCCCAAGCGCTCTTCAATTT 4353  
Qy 3065 CCTCGYACCTGTGGCTTACAGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGCAGGT 3124  
Db 4354 CCTCGCTACCTGTGGCTTACAGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGCAGGT 4413  
Qy 3125 TAAATACTTAGGRTAARATTTATCCAAAGGCCACCGGCCCTCAGTGAGGAAYATCCA 3184  
Db 4413 TAAATACTTAGGRTAARATTTATCCAAAGGCCACCGGCCCTCAGTGAGGAAYATCCA 3242

Db 4414 ----TACTTAGGGCTAAAATTTATCCAAAGGCACAGGGCCCTCAGTGAGGAACACATCCA 4469  
Qy 3185 GCTATACTGGCTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGCRT 3244  
Db 4470 GCTATACTGGCTTATCTCTCATCCAAACCCCTAAAGCAACTAAGGGGATTCCTTGGCGT 4529  
Qy 3245 AAYAGYTTCTCCGGAATATGATTCCTCCAGGTGTCGGAATATAGCCAGGYCATTAATA 3304  
Db 4530 AATAGGTTTCTCCGGAATATGATTCCTCCAGGTGTCGGAATATAGCCAGGYCATTAATA 4588  
Qy 3305 CASTAATTAAGCAAACTCAGAAAGCAATACCCATTTARTAGATGGAAMCTGAAGYMR 3364  
Db 4589 CACTAATTAAGCAAACTCAGAAAGCAATACCCATTTARTAGATGGAAMCTGAAGTAG 4648  
Qy 3365 AAGTGGCTTCTCCAGGCCCTTAAAGAGGCTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 3424  
Db 4649 AAGTGGCTTCTCCAGGCC-----TAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 4694  
Qy 3425 AACRGGGCAAGACTTTTSTTATATATATATATATATATATATATATATATATATATAT 3484  
Db 4695 AACRGGGCAAGACTTTTSTTATATATATATATATATATATATATATATATATATATAT 4753  
Qy 3485 TTACACAGTCTCCAGGAGTGTGCAACCTGTGGCATACCTGTAGGAAATATGATG 3544  
Db 4754 TTACACAGTCTCCAGGAGTGTGCAACCTGTGGCATACCTGTAGGAAATATGATG 4813  
Qy 3545 TAGTGGCAAGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTGTAGTAT 3604  
Db 4814 TAGTGGCAAGGTTGRCYTCATTGTTTAYGGGTAGTGGTGGCAGTGTAGTAT 4873  
Qy 3605 CTGAAGCAGTTAAATAAT 3664  
Db 4874 CTGAAGCAGTTAAATAAT 4933  
Qy 3665 AYRGCACTCTCATCTGCTTAAGGAGACTGTGGCTGTGCAACAACCTGTAAATATTC 3724  
Db 4934 ATGGCACTCTCATCTGCTTAAGGAGACTGTGGCTGTGCAACAACCTGTAAATATTC 4993  
Qy 3725 AGGCTCTATTACTTGAAGGCCAGTCTGCACTGTGCACTGTGCACTGTGCACTGTGCACT 3784  
Db 4994 AGGCTCTATTACTTGAAGGCCAGTCTGCGACTGTGCACTGTGCACTGTGCACTGTGCACT 5053  
Qy 3785 YNCATTTCTTCAGACAATGAAGAAAGATATATATATATATATATATATATATATAT 3844  
Db 5054 CCACATTTCTTCAGACAATGAAGAAAGATATATATATATATATATATATATATATAT 5113  
Qy 3845 ARACTTATGCCACTCCAGGGGACTTGTAGAGTTCCTTGTGACTGTGACTGTGACTGTGACT 3904  
Db 5114 ARACTTATGCCACTCCAGGGGACTTGTAGAGTTCCTTGTGACTGTGACTGTGACTGTGACT 5172  
Qy 3905 TTGTATACTGATGGAAGTTCTTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAAGT 3964  
Db 5173 TTGTATACTGATGGAAGTTCTTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAAGT 5232  
Qy 3965 GTCAGTGATAATGGAATATTTGAAAGTAATCCCTCATCTCCAGGAACCTAGTGTACGTR 4024  
Db 5233 GTCAGTGATAATGGAATATTTGAAAGTAATCCCTCATCTCCAGGAACCTAGTGTACGTR 5292  
Qy 4025 GCAGAACTAATACCCCTCATYKGGGCACTAGATTTAGGAGAAAGAAAGGGAATATA 4084  
Db 5293 GCAGAACTAATACCCCTCATYKGGGCACTAGATTTAGGAGAAAGAAAGGGAATATA 5352  
Qy 4085 TATACAGACTCTTARTATGTCTACCTAGTCTCCTCATGCCCCATGMRGCAATATGSAAGAA 4144  
Db 5353 TATACAGACTCTTARTATGTCTTACCTAGTCTCCTCATGCCCCATGSAAGAA 5412  
Qy 4145 AGGAAATTCCTAACTTCYAGAGGAACCTTATCAMACATCAGGAAGCCATTAGARATTA 4204  
Db 5413 AGGAAATTCCTAACTTCYAGAGGAACCTTATCAMACATCAGGAAGCCATTAGARATTA 5472  
Qy 4205 TTATYGGCWGTACAGAAACCTTATAGAGGTGAGTCTTACACTGCGGGGTTCATCANAAA 4264  
Db 5473 TTATYGGCWGTACAGAAACCTTATAGAGGTGAGTCTTACACTGCGGGGTTCATCANAAA 5532



QY 4265 GGAAAGGAAAGGAAATGAAAGGAAATGCAAGCAATATTTGAAGCAAAAGAGCTGCA 4324  
DB 5533 GGAAAGGAAAGGAAATGAAAGGAAATGCAAGCAATATTTGAAGCAAAAGAGCTGCA 5592  
QY 4325 AGGCAAGGACCTCCATGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTC 4384  
DB 5593 AGGCAAGGACCTCCATGAAATGCTTATTAACCAACCTTCCCTTAGTATAGGGTAAATCCCTTC 5652  
QY 4385 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTCAGGAG-CAG 4443  
DB 5653 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTCAGGAGACAG 5712  
QY 4444 TTTTCTCCCTCGGAGCGTTAGCCACTGAGAGGAAATATCTTTTCCCTGCAACTAT 4503  
DB 5713 TTTTCTCCCTCGGAGCGTTAGCCACTGAGAGGAAATATCTTTTCCCTGCAACTAT 5772  
QY 4504 CCAATGGAAATTAACCTTAAACCTTCATCAACCTTTGCTAGGATCGATAGCACCA 4563  
DB 5773 CCAATGGAAATTAACCTTAAACCTTCATCAACCTTTGCTAGGATCGATAGCACCA 5832  
QY 4564 TCARATGGCAATCAATTAATTTACTTGGACCGGCTTTTCAAACTATCAAGCABATAT 4623  
DB 5833 TCAGATGGCAATCAATTAATTTACTTGGACCGGCTTTTCAAACTATCAAGCABATAT 5892  
QY 4624 CAGGCTCTGTGAATGTGCCARAAATTAATCCCTGCTTATCGCCAAAGCTTCTTCAGG 4683  
DB 5893 CAGGCTCTGTGAATGTGCCARAAATTAATCCCTGCTTATCGCCAAAGCTTCTTCAGG 5952  
QY 4684 ABAACAAABAACAGGCAATACCTTGRABARACTGGGAACTGATTTTACCCAAAGGCC 4743  
DB 5953 AGAACAAABAACAGGCAATACCTTGGAGAACTGGCAACTGATTTTACCCAAAGGCC 6012  
QY 4744 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTABATACTTTCACGGTGGGACAG 4803  
DB 6013 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTABATACTTTCACGGTGGGACAG 6072  
QY 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCAATGAAATA 4863  
DB 6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCAATGAAATA 6132  
QY 4864 ATCCCAAGATTCGGACTTCCCGAGGCTTACAGATGACATAGCCCTGCTTTCAGGCC 4923  
DB 6133 ATCCCAAGATTCGGACTTCCCGAGGCTTACAGATGACATAGCCCTGCTTTCAGGCC 6192  
QY 4924 ACAGTAACCCAGGAGTATCCAGGCTTTAGTATACCATATCACTTACACTGGCCCTGA 4983  
DB 6193 ACAGTAACCCAGGAGTATCCAGGCTTTAGTATACCATATCACTTACACTGGCCCTGA 6252  
QY 4984 AGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAAAYACTCAAAGGACATCTAAA 5043  
DB 6253 AGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAAAYACTCAAAGGACATCTAAA 6312  
QY 5044 AAGCAAAACCCAGGAAACCCACTGCATGCGCTGTCTGTGGCTTATAGCCCTTAAAAGA 5103  
DB 6313 AAGCAAAACCCAGGAAACCCACTGCATGCGCTGTCTGTGGCTTATAGCCCTTAAAAGA 6372  
QY 5104 ATCTGCAACTTTCCCAAAAGAGGAGCTTACCCATACGAAATGCTGTATGGAGGCC 5163  
DB 6373 ATCTGCAACTTTCCCAAAAGAGGAGCTTACCCATACGAAATGCTGTATGGAGGCC 6432  
QY 5164 TTCTAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGGACACATCACCT 5223  
DB 6433 TTCTAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGGACACATCACCT 6492  
QY 5224 CCTTAGCCAAATATCAAGCTTCTTAAAACATTTCAAGGAACCTATCCCTGAGAAGG 5283  
DB 6493 CCTTAGCCAAATATCAAGCTTCTTAAAACATTTCAAGGAACCTATCCCTGAGAAGG 6552  
QY 5284 GAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCCTTAATTC 5343  
DB 6553 GAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCCTTAATTC 6612

QY 5344 CATCCTAGATACATCTCTGGGAAGGACCTACCCAGTCAATTTATYATCCCAACTGCGG 5403  
DB 6613 CATCCTAGATACATCTCTGGGAAGGACCTACCCAGTCAATTTATCTACCCCAACTGCGG 6672  
QY 5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACAACCTTGAGTCAAACTCTGGATCTGC 5463  
DB 6673 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACAACCTTGAGTCAAACTCTGGATCTGC 6732  
QY 5464 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTCTGTGAAACCTCTAGAGATT 5523  
DB 6733 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTCTGTGAAACCTCTAGAGATT 6792  
QY 5524 TGCCTCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATAAAT-CCCATG 5583  
DB 6793 TGCCTCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATAAAT-CCCATG 6851  
QY 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTTTCACTCTCTCACTGCA 5643  
DB 6852 GSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTTTCACTCTCTCACTGCA 6911  
QY 5644 CCCCCTCATGCGCTGTATGACCAAGTAGCTCCCTCTYACCMAGAGTTTCTATGGAGATTG 5703  
DB 6912 CCCCCTCATGCGCTGTATGACCAAGTAGCTCCCTCTYACCMAGAGTTTCTATGGAGATTG 6971  
QY 5704 CAGGCTCCCGAAATATTTGATGCCCATCGTATAGAGTCTTTTAAAGGGAACCCCACT 5763  
DB 6972 CAGGCTCCCGAAATATTTGATGCCCATCGTATAGAGTCTTTTAAAGGGAACCCCACT 7031  
QY 5764 TTCACTGCCACACCCCATATGTCGCGCAACTGCTATCACTCTGCCACTCTTTGCACTGCAT 5823  
DB 7032 TTCACTGCCACACCCCATATGTCGCGCAACTGCTATCACTCTGCCACTCTTTGCACTGCAT 7091  
QY 5824 GCAATACTCATTTATGACAGGAAATGATTAATCTTAGTTGCTGAGAGACTTGA 5883  
DB 7092 GCAATACTCATTTATGACAGGAAATGATTAATCTTAGTTGCTGAGAGACTTGA 7151  
QY 5884 GTCACTGCTCTGGACTTACTTCAACCAACTGTATGTCTGATGGGGGTGGAGTTCAA 5943  
DB 7152 GTCACTGCTCTGGACTTACTTCAACCAACTGTATGTCTGATGGGGGTGGAGTTCAA 7211  
QY 5944 GATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCAACTCACCGGGGTACATGCG 6003  
DB 7212 GATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCAACTCACCGGGGTACATGCG 7271  
QY 6004 ACCTTAGCCCTCTCAAGAGACTAGTCTCTCAAACTACATGAAACCCCTCGTACCCAT 6063  
DB 7272 ACCTTAGCCCTCTCAAGAGACTAGTCTCTCAAACTACATGAAACCCCTCGTACCCAT 7331  
QY 6064 ACTGCTGTTAGGCTTATTTAATACCACTCTCACTGGGCTCACTGAGGCTCTCGGCCCAA 6123  
DB 7332 ACTGCTGTTAGGCTTATTTAATACCACTCTCACTGGGCTCACTGAGGCTCTCGGCCCAA 7391  
QY 6124 AACCTACTAACTGTTGGATATGCTCCCTGAACTTTCARGGCTATGTTTCAATCCCT 6183  
DB 7392 AACCTACTAACTGTTGGATATGCTCCCTGAACTTTCARGGCTATGTTTCAATCCCT 7451  
QY 6184 GTACTGAACTAGGAACTTTCAGCAGCAAGAAATAAACACCACTTCCGTTTTAGTAGA 6243  
DB 7452 GTACTGAACTAGGAACTTTCAGCAGCAAGAAATAAACACCACTTCCGTTTTAGTAGA 7511  
QY 6244 CCTCTGTTTCCAAATGGAATAAACCATACCTCAAACTCACTGTTGTAATAATTTAGC 6303  
DB 7512 CCTCTGTTTCCAAATGGAATAAACCATACCTCAAACTCACTGTTGTAATAATTTAGC 7571  
QY 6304 AATACTACATACCAACCACTCCCAATGCACTAGGTTGGGTAACTCTCTCCCAACCAATA 6363  
DB 7572 AATACTACATACCAACCACTCCCAATGCACTAGGTTGGGTAACTCTCTCCCAACCAATA 7631  
QY 6364 GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATGCTGTTTGAAT 6423  
DB 7632 GTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATGCTGTTTGAAT 7691  
QY 6424 GGCTCTTTCAGAACTATGCT 6483





```

4865 TTCCCGAGATCGGACTTCCCGGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCCA 4924
16911 TTCCCGAGATCGGACTTCCCGGAGGCTTACAGAGTGACAAATAGCCCGCTTTCAAGGCTG 16852
4925 CAGTACCCAGGAGGATCCCGGCG- - -TTAGGTATACGATCACTACCTACCTGCGCC 4980
16851 CAGTACCCAGGAGGATCGCAGCGTTACTTAGGCATCAATATCACTTACACTGCGCC 16792
4981 TGAAGGCCACAGTCCCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAAAGGACATCTA 5040
16791 TGGAGGCCACATCTCTCAGGAGAGTTCGAGAAATGAACGAACACTCAACACATCTA 16732
5041 AAAAGCAACCCAGAAACCCACCTCACATGCGCTGTCCTGCTATAGCCTTAAAA 5100
16731 AAAAGCTAACTCAGGAAACCCACCTCGCATGCGCTGCTGTTGCCATAGCCTTATTA 16672
5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAG 5160
16671 AGAATACAAAACTCCCGCCCAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAG 16612
5161 CCCTTCATACCAATGACCTTGCTTGACCCA- AGACAGGCCAACTTAGTTGCAGACATC 5219
16611 CCCTTCATACCAATGACCTTGCTTGACCCAAGGAGAGCGGCCAATTTAGTTGCAGACATC 16552
5220 ACTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAAGGAACCTATCCCTGAGAA 5279
16551 ACTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAAGGAACCTGTATCCCGAGAG 16492
5280 GAGGGAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAT 5339
16491 GAGGGAAGGA- - -ATTCCACCCCTGGTAAATGATAGTCAAGTCCCTTCCTCTAAG 16435
5340 TCCCATCTCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATVATCCCAACT 5399
16434 TCCCAACCCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATVATCCCTGAGT 16375
5400 GCGGTTAAAGTGGCTGGAGTCTTTGGATACATCAACATTTGAGTCAAAATCCTGGATA 5459
16374 GCAAGTTAAAGTGTCTGGAGTGGAGTCTTAGACACA- - - - -TCMAACCCCTGGATA 16326
5460 CTGCGAAGGAACCTGAAATTCAGGAGACAAACGCTAGCTATTCTGTGACCTCTAGAG 5519
16325 CGCCAAAGGAACCCAAAATTCAGGAGACAAACGCTAGCTATTCTGTGAACCTCTAGAG 16266
5520 GATTTCGCGCTGCTTCAACCAACCAACAGGAGGAAGTAACATAAAATCATAAATCCCC 5579
16265 GATCTGACCTGCTCTTCAAGCAGACAACTCGAGGAAGTAAGTAAGTGAATCTGAGAGTCC 16206
5580 CATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTA- CCCTSTTTCACTCTCA 5638
16205 - -ATGGCCCCCCCCGTCATATTTTCTCTTACTGTGTTGCTTACCCCTTTTCACTATCA 16148
5639 CTGACCCCTCCATGCGCTGTATGACCAAGTACGCTCCCTCTYACCMAGAGTTTCTATGGA 5698
16147 CCTCACCCCTCCATGCGCTATACCTACCAAGTACGCTCCCTTACCAAGAGCTTCTATGGA 16088
5699 GANTCAGCGTCCCGGAATATTGATGCCCATCTGATAGAGTCTTTTAAAGGGAACCC 5758
16087 GAATGCGGCTTCCGAGAAATATTGATGCCCATCTGATAGAGTCTTTTAAAGGGAACCC 16028
5759 CCACTTCACTGCCACACCCATATGCCCCGCAACTGC 5796
16027 CCACTTTCACATCCACACCCATATGCCCCGCAACTGC 15990
```

RESULT 4  
US-09-949-016-12249/c  
; Sequence 12249, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249
```

```

Query Match 39.5%; Score 2994.2; DB 4; Length 77997;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

QY 1812 ACTTTGGAGGCTCTGGAAAGGGAAGCTGGGCAAAATGAAATGCCCTAATAGGCTTGCT 1871
DB 19935 ACTTTGGAGGCTCTGGAAAGGGAACGCTGGGCAAAATCAAATGCCGTATAGGCTTGCT 19876
QY 1872 TCCAGTCCGCTACAGAGACACTTTAAAAAGATGTTCCNAGTAGAGTAAGCCGCC 1931
DB 19875 TCCAGTGTGCTCTACAAGGACACTTTAAAAAAGATGTTCTGAATAGAAATTAAGCTGCC 19816
QY 1932 TTGCTCATGCCCTTATTTTCAAGGAATCACTGGAAGGCCCACTGCCAGGGGACAAA 1991
DB 19815 -TCATCATGCCCTCATGTCAAGGGAATCACTGGAAGGCCCACTGCCAGGGGATGAA 19757
QY 1992 GGTCTTTTGAAGTCAAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGCTGCTGG 2051
DB 19756 GGTCTTTTCAAGTCAAGAGCCACTAACCCAGATGATGAGGCTGAGGCTGCTGG 19697
QY 2052 GGCAGCGGCATCCCATGCCATCAACCTCACAGAGCCCTGGGCTATGTTGACCATTTGAGG 2111
DB 19696 GGCAGCGGCAGCCCAT- - - - -GCCCCAGGCTATGCTTGACCATTTGGG 19654
QY 2112 GCCAGGAAGGT- - -TGCTCTCTGGACACTGGTGGCTTCTTCTAGTCTTACTTCTCTGTC 2168
DB 19653 GCCAGGAAGGTAACTGTCTCTGAACTGGCATGGCCCTTCTCAGTCTTAGTCTCTGCTC 19594
QY 2169 CCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGCTCCNTAAGACGGGCACTCA 2228
DB 19593 CTGGACAAACGCTCTCCAGATCTGTCACTACCC- - -GAGGGGCTCTAGGATAGGCACTCA 19537
QY 2229 CTAGATACCTTTTTCAGGCCCACTAAGTTATGAATCTGGGAGCTTTATTCTTTTCACTATGC 2288
DB 19536 CTAGATAC- - -TTCTCCAGGCCCACTAAGTTGTG- - -ACTGGGGAACCTTCACTTTTTCAC 19479
QY 2289 TTTTCTTAATTTATGCTTTGAAAGCCCACTACTGTTTAGGGAGAGACATTTCTAGCAAAAGC 2348
DB 19478 CTTTCTTAATTTATGCTTTGAAACCCCACTACTGTTTAGGGAGATACATTTCTAGCAAAAGC 19419
QY 2349 AGGGGCCATTTATACACTTGAACTAGGAGAAAGCAACCCGCTTTGTTGTTGTTGTTGTTGTTG 2408
DB 19418 AGGGGCCATTTATACACTTGAACTAGGAGAAAGCAACCCGCTTTGTTGTTGTTGTTGTTGTTG 19361
QY 2409 AGGAAGGAATTAATCTGGAAGTCTGGGCAACAGAGGAACATATGACGAGCCCAAGAAAT 2468
DB 19360 AGGAAGGAATTAATCTGGAAGTCTGGGCAACAGAGGAACATATGATGATGAG- - -GGAAAGAT 19302
QY 2469 GCCCGCTCTGTTCAAGTTAACTAAAGGATTCACCTTCCCTTCCCTACCAAGGACGATAC 2528
DB 19301 GCCCGTCCCGCTTCAAGTTAACTAAAGGATTCGCGCTCTCTTTTCCCAACCAAGGACGATAC 19242
```



Db 17095 AACAAAGAAAGGCGCAATACCCAGGAGAGACTGGCAACTAGATTTTACCCACATGCCCA 17036  
Qy 4745 AACCTCAGGATTTTCAGTATCTACTAGCTGGGTATATATCTTTCAGGGTTGGGCARAGG 4804  
Db 17035 AATCTCAGGATTTTCAGTATCTACTAGCTGGATAGATATCTTTCAGTGGTTGGAGAGG 16976  
Qy 4805 CCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTATATAAGGCACTAGTTTCATGAATAA 4864  
Db 16975 CCTTCTCTAGGACAGAAAGGCCCAAGAGGTATATAAGGCACTAATTCATGAATAA 16916  
Qy 4865 TTCCAGATTCGAGTTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCAGGCCA 4924  
Db 16915 TTCCAGATTCGAGTTTCCCTGAGACTTACAGAGTGAACAATAGCCCTGCTTTCAGGCCG 16856  
Qy 4925 CAGTACCCAGGAGTATCCAGGCG---TTAGGTATAGATATCACTTACACTGGCC 4980  
Db 16855 CAGTACCCAGGAGTATCCAGGCGTTACTTAGGCATACAATATCACTTACACTGGCC 16796  
Qy 4981 TGAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAATCAATGAAAGGACATCTA 5040  
Db 16795 TGAAGGCCACATCTCAGGGAAGTTCGAGAAATGAATGAAAGGACATCTAATCAATCTA 16736  
Qy 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGCTGCTGCTGCTTATAGCTTAA 5100  
Db 16735 AAAAAGCTAACTCAGGAAACCCACCTCGCATGGCTGCTGCTGCTGCTGCTTATTA 16676  
Qy 5101 AGAATCTGAACCTTCCCAAAAGGAGGACTTACCCATACGAAATGCTGATGGAAG 5160  
Db 16675 AGAATCAAAACTCCCAAAAGGAGGACTTACCCATACGAAATGCTGATGGAAG 16616  
Qy 5161 CCCTTCATACCAATGACTTGTCTGTGACCCA-AGACAGCAACTTATGTTGAGACATC 5219  
Db 16615 CCCTTCATACCAATGACTTGTCTGTGACCGAGAGCGCAATTTAGTTGAGACATC 16556  
Qy 5220 ACTCTCTAGCAAAATCAACAAGTTCTTAAACATTTACAAGAACTATCCCTGAGAA 5279  
Db 16555 ACTCTCTAGCAAAATCAACAAGTTCTTAAACATTTACAAGGAACTGTACCCGAGAG 16496  
Qy 5280 GAGGGAAGAACTATTTCCACCCGAGGAGTATGATGATGATGATGATGATGATGATGAT 5339  
Db 16495 GAGGGAAGGA---ATTCCACCTGGTAACTGATGATGATGATGATGATGATGATGAT 16439  
Qy 5340 TCCCATCTCTAGATACATCTGGAAGGACCTTACCAGTCATTTTATATATACCCCACT 5399  
Db 16438 TCCCATCTCTAGATACATCTGGAAGGACCTTACCAGTCATTTTATATATCTACCTGAT 16379  
Qy 5400 GCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTCAGTCAAAATCTGGATA 5459  
Db 16378 GCAGTTAAAGTGTCTGGAGTGGAGTCTTGGACACA-----TCNAACCCCTGGATA 16330  
Qy 5460 CTGCAAAAGGAACCTGAAATTCAGGAGACAACGCTAGCTATTCTGTGAACTCTAGAG 5519  
Db 16329 CCGCAAAAGGAACCCCAAAATTCAGGAGACAACGCTAGCTATTCTGTGAACTCTAGAG 16270  
Qy 5520 GATTTGGCTGCTCTTCARACACAACACGAGGAAAGTAACTAAATCATATAATCCCC 5579  
Db 16269 GATCTGACCTGCTCTTCAGGAGACAACGAGGAAAGTAACTGGAATCTGTAGCTCC 16210  
Qy 5580 CATGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTAA-CCCTTTTCTACTCTCA 5638  
Db 16209 --ATGGCCCCCGCTCATATTTTCTCTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 16152  
Qy 5639 CTGACCCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGA 5698  
Db 16151 CCTCACCCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGA 16092  
Qy 5699 GAATGACGCTCCCGGAATATGATGCCCCCATCTGATAGAGTCTTTTSTAAGGGAACCC 5758  
Db 16091 GAATGCGGCTTCCCGAATATGATGCCCCCATCTGATAGAGTCTTTTSTAAGGGAACCC 5758  
Qy 5759 CCACCTTCACTGCCCCACACCCCATATGCCCCGCACTGC 5796

Db 16031 CCACCTTCCATCCACACCCCATATGCCCCGTGCACCTTC 15994  
RESULT 5  
US-09-175-928-3  
; Sequence 3, Application US/09175928A  
; Patent No. 6312921  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallie, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Mi, Sha  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: 6006B.AU172A  
; CURRENT APPLICATION NUMBER: US/09/175,928A  
; CURRENT FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2946  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-175-928-3  
Query Match 38.1%; Score 2888.6; DB 3; Length 2946;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 2872; Conservative 49; Mismatches 3; Indels 1; Gaps 1;  
Qy 4658 CTCCTTATCGCCCAAGCTCTCTCAGGAAACAAARAAACAGGCCATTACCTGRRARAC 4717  
Db 6 CTCCTTATCGCCCAAGCTCTCTCAGGAAACAAARAAACAGGCCATTACCTGGAAGAC 65  
Qy 4718 TGGCAACTGATTTTACCACAAGCCCAACCTCAGGGATTTAGTATCTACTAGCTGGG 4777  
Db 66 TGGCAACTGATTTTACCACAAGCCCAACCTCAGGGATTTAGTATCTACTAGCTGGG 125  
Qy 4778 TARATCTTTCAGGGTTGGGCARAGGCTTCCCTCTAGGACAGAAAGGCCCAAGAGG 4837  
Db 126 TAGATATCTTTCAGGGTTGGGCARAGGCTTCCCTCTAGGACAGAAAGGCCCAAGAGG 185  
Qy 4838 TAATAAAGGCACATGTTTCATGAATAATTCAGATTCGGACTTCCCGAGGCTTACAGA 4897  
Db 186 TAATAAAGGCACATGTTTCATGAATAATTCAGATTCGGACTTCCCGAGGCTTACAGA 245  
Qy 4898 GTGCAATAGCCCTGCTTTCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTA 4957  
Db 246 GTGCAATAGCCCTGCTTTCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTA 305  
Qy 4958 TACGATATCTTACATCTGCGCTGGAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGA 5017  
Db 306 TACGATATCTTACATCTGCGCTGGAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGA 365  
Qy 5018 ATGAAAYACTCAAGGACATCTTAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTG 5077  
Db 366 ATGAAAYACTCAAGGACATCTTAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTG 425  
Qy 5078 YTCGTTGCTTATAGCTTAAAAAGAAATCTGCAACTTTTCCCAAAAGAGGACGACTTAGCC 5137  
Db 426 CTCGTTGCTTATAGCTTAAAAAGAAATCTGCAACTTTTCCCAAAAGAGGACGACTTAGCC 485  
Qy 5138 CATACGAAATGCTGTATGGAAGGCCCTTCTATACCAATGACCTTGTGCTGACCCCAAGC 5197  
Db 486 CATACGAAATGCTGTATGGAAGGCCCTTCTATACCAATGACCTTGTGCTGACCCCAAGC 545  
Qy 5198 AGCCAACTTAGTTCGAGACATCCTCTCTAGCCAAATATCAACAAGTTCTTAAACATT 5257  
Db 546 AGCCAACTTAGTTCGAGACATCCTCTCTAGCCAAATATCAACAAGTTCTTAAACATT 605



Qy	5258	ACAAGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCCACCWGTGACATGGTATT	5317
Db	606	ACAAGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCCACCCTTGTGACATGGTATT	665
Qy	5318	AGTCAAGTCCCTTCTCTAAATTCGCCATCCCTAGATACATCCTCGGGAAGGACCCCTAGCC	5377
Db	666	AGTCAAGTCCCTTCTCTAAATTCGCCATCCCTAGATACATCCTCGGGAAGGACCCCTAGCC	725
Qy	5378	AGTCATTATTATYACCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA	5437
Db	726	AGTCATTATTATYACCCAACTGCGGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCA	785
Qy	5438	CACCTGAGTCAAACTCCTGGATATCTGCCAAAGGAACCTGAAAATCCAGGAGCAACGCTAG	5497
Db	786	CACCTGAGTCAAACTCCTGGATATCTGCCAAAGGAACCTGAAAATCCAGGAGCAACGCTAG	845
Qy	5498	CTATTCCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAAACCGAGGAGAAA	5557
Db	846	CTATTCCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAAACCGAGGAGAAA	905
Qy	5558	GTAACATAAAATCATAAAATCCCCATGSGSCCTCCCTTATCATATTTTCTCTKTASTGTT	5617
Db	906	GTAACATAAAATCATAAAATCCCCATGSGSCCTCCCTTATCATATTTTCTCTKTASTGTT	964
Qy	5618	TTTTACCCCTSTTTCACTCTCACTGCAACCCCTCCATGCGGCTGTATGACCACTAGTCTCC	5677
Db	965	TTTTACCCCTSTTTCACTCTCACTGCAACCCCTCCATGCGGCTGTATGACCACTAGTCTCC	1024
Qy	5678	CTYACCMAGAGTTTCTATGGAAGATGACAGCGTCCCGGAAATATTGATGCCCCATCGTATA	5737
Db	1025	CTYACCMAGAGTTTCTATGGAAGATGACAGCGTCCCGGAAATATTGATGCCCCATCGTATA	1084
Qy	5738	GGAGTCTTTTAAAGGGAACCCCACTTCACTGCCCCCAACCCATATGCGCGCACTGCT	5797
Db	1085	GGAGTCTTTTAAAGGGAACCCCACTTCACTGCCCCCAACCCATATGCGCGCACTGCT	1144
Qy	5798	ATCACTCTGCCACTCTTTGATGATGCAAAATFACTATTATGGACAGGAAAAATGATTA	5857
Db	1145	ATCACTCTGCCACTCTTTGATGATGCAAAATFACTATTATGGACAGGAAAAATGATTA	1204
Qy	5858	ATCCTAGTTGCTGGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCAACCCAACTG	5917
Db	1205	ATCCTAGTTGCTGGAGGACTTGGAGTCACTGCTGTGTGGACTTACTTCAACCCAACTG	1264
Qy	5918	GTATGCTGTAGTGGGGTGGAGTTCAAGATCAGGCAAGGAAACATCTNAAGAGTAA	5977
Db	1265	GTATGCTGTAGTGGGGTGGAGTTCAAGATCAGGCAAGGAAACATCTNAAGAGTAA	1324
Qy	5978	TCTCCCAACTCACCGGGGTACATGGCACTCTAGCCCCCTACAAAGGACTAGATCTCTCAA	6037
Db	1325	TCTCCCAACTCACCGGGGTACATGGCACTCTAGCCCCCTACAAAGGACTAGATCTCTCAA	1384
Qy	6038	AACTACATGAACCCCTCGGTACCCATATCTGCGTGGTAAAGCTATTAAATACCACTCTCA	6097
Db	1385	AACTACATGAACCCCTCGGTACCCATATCTGCGTGGTAAAGCTATTAAATACCACTCTCA	1444
Qy	6098	CTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTTAACTGTGGATATGCTCCGCCCTGA	6157
Db	1445	CTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTTAACTGTGGATATGCTCCGCCCTGA	1504
Qy	6158	ACTTCARGCCATATGTTTCAATCCCTGTACCTGAAACAAATGGAACAACTTCAGCACAGAAA	6217
Db	1505	ACTTCARGCCATATGTTTCAATCCCTGTACCTGAAACAAATGGAACAACTTCAGCACAGAAA	1564
Qy	6218	TAAACACACCTTCGGTTTTAGTAGGACTCTGTGTTTCCAAATSTGGAATAAACCCATACCT	6277
Db	1565	TAAACACACCTTCGGTTTTAGTAGGACTCTGTGTTTCCAAATSTGGAATAAACCCATACCT	1624
Qy	6278	CAAACTCCTGCTGTAAATTTAGCAATACCTACATACACCAACCACTCCCAATGCAATCA	6337
Db	1625	CAAACTCCTGCTGTAAATTTAGCAATACCTACATACACCAACCACTCCCAATGCAATCA	1684
Qy	6338	GGTGGTAACTCTCCCAACAAATAGTCTGCTACCTCCCTCAGGAATATTTTTTGTCTGTG	6397

Db	1685	GGTGGTAACTCTCCCAACAAATAGTCTGCTACCCCTCAGGAATATTTTTGTCTGTG	1744
Qy	6398	GTACTCTAGCCTATTCGTTGTTTGAATGGCTCTTTCAGAAATCTATGCTCTCTCTCATTTCT	6457
Db	1745	GTACTCTAGCCTATTCGTTGTTTGAATGGCTCTTTCAGAAATCTATGCTCTCTCTCATTTCT	1804
Qy	6458	TAGTGCCCCCTATGRCATCTACACTGAACAAAGATTTTATACAGTTATGTCATATCTAAGC	6517
Db	1805	TAGTGCCCCCTATGRCATCTACACTGAACAAAGATTTTATACAGTTATGTCATATCTAAGC	1864
Qy	6518	CCCGCAACAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTGTAGTGGTGCAC	6577
Db	1865	CCCGCAACAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTGTAGTGGTGCAC	1924
Qy	6578	TAGTACTTGGGATTTGGCGGTATCAACCTCTACTCAGTTCTACTACAAAATCTATCTAAG	6637
Db	1925	TAGTACTTGGGATTTGGCGGTATCAACCTCTACTCAGTTCTACTACAAAATCTATCTAAG	1984
Qy	6638	AACTAAATGGGACATGGAACGGGTCCCGGACTCCCTGCTCACCTTGCAGAGATCAACTTA	6697
Db	1985	AACTAAATGGGACATGGAACGGGTCCCGGACTCCCTGCTCACCTTGCAGAGATCAACTTA	2044
Qy	6698	ACTCCCTAGCAGCAGTACTCTTCRAAATCGAAGAGCTTTTAGACTGTCTAACCGCTGARA	6757
Db	2045	ACTCCCTAGCAGCAGTACTCTTCRAAATCGAAGAGCTTTTAGACTGTCTAACCGCTGARA	2104
Qy	6758	GAGGGGAACTCTGTTTATTTTATAGGGGAAGATGCTGTTTATATGTTTAAATCAATCCGAA	6817
Db	2105	GAGGGGAACTCTGTTTATTTTATAGGGGAAGATGCTGTTTATATGTTTAAATCAATCCGAA	2164
Qy	6818	TCGTCACCTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGATAGCAGAGGAGTTC	6877
Db	2165	TCGTCACCTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGATAGCAGAGGAGTTC	2224
Qy	6878	GAACACCTGGACCTCGGGGCTCTCAGCCCATGGATGATCCCTGGATTTCTCCCTCTCTTAG	6937
Db	2225	GAACACCTGGACCTCGGGGCTCTCAGCCCATGGATGATCCCTGGATTTCTCCCTCTCTTAG	2284
Qy	6938	GACCTCTAGCAGCTATTAATATGCTACTCTCTTTGGACCCCTGTATCTTTTACCTCTCTTG	6997
Db	2285	GACCTCTAGCAGCTATTAATATGCTACTCTCTTTGGACCCCTGTATCTTTTACCTCTCTTG	2344
Qy	6998	TTAATTTGCTCTCTTCAGAAATCGAAGCTGAAATACAAATGAGCCCAAGATGCACT	7057
Db	2345	TTAATTTGCTCTCTTCAGAAATCGAAGCTGAAATACAAATGAGCCCAAGATGCACT	2404
Qy	7058	CCAAAGACTTAAGATCTACCGCAGACCCCTGGACCCGCTGTAGCCACGATCTGATGTTA	7117
Db	2405	CCAAAGACTTAAGATCTACCGCAGACCCCTGGACCCGCTGTAGCCACGATCTGATGTTA	2464
Qy	7118	ATGACATCAAAAGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCA	7177
Db	2465	ATGACATCAAAAGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCA	2524
Qy	7178	ATTACAGAGGAGCAGTTAGAGCGGTGCTGGCCAACTCCCTCCCAACAGCAGCTTAGGTTT	7237
Db	2525	ATTACAGAGGAGCAGTTAGAGCGGTGCTGGCCAACTCCCTCCCAACAGCAGCTTAGGTTT	2584
Qy	7238	CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAT	7297
Db	2585	CCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAT	2644
Qy	7298	CCYTAAGCCTAGTGGGAGGTGACCATCTCCCTTTAAACACAGGGGCTTGCAACTTAG	7357
Db	2645	CCCTAAGCCTAGTGGGAGGTGACCATCTCCCTTTAAACACAGGGGCTTGCAACTTAG	2704
Qy	7358	YTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAAATAGGCAAGACAGGAGTAA	7417
Db	2705	CTCACCTGACCAATCAGAGAGCTCACTAAATGCTAAATAGGCAAGACAGGAGTAA	2764
Qy	7418	AGAAATAGCAATCATTTATGCTGAGAGCAGCAGGAGGACCAATGATCGGGATATA	7477

```
Db 2765 AGAATAGCCCAATCATCTATTGCTGAGAGCACAGCAGGAGGACAATGATCGGGATATA 2824
Qy 7478 AACCCAAAGTTCGAGCGCGCAACGCCCTTTGGGTCCCTCCCTTTGTTGATGGGA 7537
Db 2825 AACCCAAAGTTCGAGCGCGCAACGCCCTTTGGGTCCCTCCCTTTGTTGATGGGA 2884
Qy 7538 GCTCTGTTTTCATGCTATTTTCACCTCTATTAAATCTTGCAATGCTGCR 7582
Db 2885 GCTCTGTTTTCATGCTATTTTCACCTCTATTAAATCTTGCAATGCTGCA 2929

RESULT 6
US-09-949-016-13002/c
; Sequence 13002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13002
; LENGTH: 168394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(168394)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-13002

Query Match 34.5%; Score 2615.4; DB 4; Length 168394;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 3133; Conservative 134; Mismatches 355; Indels 58; Gaps 33;

Qy 1012 TTCCCCCAATATGACCCCTCCCAAGCAGTGGGAGGAGAGATTCGGCCCAAGCCAGAGTG 1071
Db 29005 TCCCCCAAGGCAAAACACCCCTAAATGTATTCTGGAGAAATTCGGCCCTGTGAGAGTG 28946
Qy 1072 CATGTGCTTTTYYTCTCCAGACTTAAAGCAAAATAAAACAGAGCTTAGGTAAATTTCTCA 1131
Db 28945 TATGTACCTTTTCCCTGTGACACTTGAAGCAAAATAAAATAGACCTTAGGTAAATTTCTCA 28886
Qy 1132 GATAAATCTGATGGCTATTTGTTGTTTAAAGGGTTAGACAAATCTTTGATCTGACA 1191
Db 28885 GATAAATCTGTTGGCTATTTGATGCTTTTAAAGGGTTAGACAAATCTTTGATCTGACA 28826
Qy 1192 TGGAGAGATATATATGTCACCTGCTTAATCAGACACTAACCCCAATGAGAGAGTGGCAC 1251
Db 28825 TGGAGAGATATA-ATGTTACTGCTAGATCAGACACTAATCCCAATGAGAGAGTGGTC 28767
Qy 1252 CATAACTCAGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGG 1311
Db 28766 CATAACTCAGCCTGAGAGTTTGGTATCTCTGGTATCTCAGTCAAGTCAATG--ATAGG 28709
Qy 1312 ATGACAAACAGAAAGAAAGANAATGATCCCAAGCCAGCAGCAGTTCACAGTCTASA 1371
Db 28708 ATGACAAACAG-AGGAAAAGAACAAATTCACAGCCAGCAGCAGTTCACAGTCTASA 28650
Qy 1372 CCCTCATTTGGGACACAGAAATCAGTAACATGGGAGATTTGGTCTGCGAGACATTTGCTAA 1431
Db 28649 CCCTCATTTGGG-----ACACAGAAATCAGAAATGGAATTTGGTGGCCACAGACATTTGCTAA 28595
Qy 1432 CTGTGTGCTASAGGAGCTAAGGAAACATASGAAGAAATCTATYGAATTAATCAATGATG 1491
```

```
Db 28594 CTTGCTTGCTA-AAGGAACTAAGGAAACTAGGAAG-AAGCCTATAGATTATTCAGTATG 28537
Qy 1492 TCACCATACACAGAGGGAGGGAAGAAATCCTACTGCTCTTCTGAGAGACTAAGGGA 1551
Db 28536 TTCACTATAACACA-AGGAAAGGAAGAAATCCTACTGCTCTTCTGAGAGACTAAGGGA 28478
Qy 1552 GGCATTGAGGAAGCGTCTCTGTGCTCCTGCTGCTCTTCTGAGAGGCAACTAATCTTTAAA 1611
Db 28477 GGCATTGAGGAAGC-AAACTCTCTGTGCTGCTGCTGTTGAGGCACTAATCTTTAAA 28419
Qy 1612 GCCTAAATTTATCATCTCAGTCAGTCAGACATTTAG-AAAAAATCTCAAAAGTCTGCCGT 1670
Db 28418 GGTTAAGTTTATCACTCAGTCAGTCAGACATTTAGAAAAAACTTTTAAAGTCCACCT 28359
Qy 1671 AGGCCGGAGCAAACTTTAGAAACCTTATTTGAACTTGGCAACCTTCGGTTTTTTATAATAG 1730
Db 28358 AGGCCAGAGCAAACTTTAGAAACCTTATTTGAACTTGGCAACCTTCGGTTTTTTATAATAG 28299
Qy 1731 AGATCAGGAGGAGCAGCGGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAG 1790
Db 28298 AGATCAGGAGGAGCAGCGGGAACAGGATTAACGGGA-TAAAGAAAGGCCACCACTTTAG 28240
Qy 1791 TCATGACCCCTCAGCAGAG-TGGACTTTTGGAGGCTCTGGAAGGAAAGCTGGGCAAT 1849
Db 28239 TCATGGCCCTCAGCAGACAGACTTCGAGGGCTCTGGAACAGGAAAGCTGGGCAAT 28180
Qy 1850 TGAATGCTTAATAGGCTTCCTCCAGTCGGTCTACAAGGACACTTTTAAAAAGATTTGT 1909
Db 28179 TGAATGCTTAATAGGCTTCCTCCAGTCGGATCTACAAGGACACTTTTAAAAAGATTTGT 28120
Qy 1910 CCAAGTAGAAGTAAGCC-GCCCTTTCCTCATGCTCCCTTATTTCAAGGGAATCACTGGAA 1968
Db 28119 TCTAATAGAAATTAAGCAACCCCTCTCATGCTTATGTATCAAGGGAATCACTGGAA 28060
Qy 1969 GGCCCACTGCCCAAGGGGACAAAGTCTTTGAGTCAGAGGCCACTAACCGAGATGATCCA 2028
Db 28059 GGCCCACTGCCCAAGGGGACGAAGTCTCTGAGTCAGAGGCCACTAACCGAGATGATCCA 28000
Qy 2029 GCAGCAGACTGAGGTCGCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCC 2088
Db 27999 GCAGCAGACTGAGGTCGCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCC 27940
Qy 2089 CTGGGTATGCTTGACCACTTGAAGGCCAGGAGGT---TGTCTCTGGAACACTGGTGGCGT 2145
Db 27939 CCGGTATGCTTAACCACTTGAAGGCCAGGAGGTAACTGTCTCTGGAACACTGGCGTGGC 27880
Qy 2146 CTTCTTATGCTTACTCTTCTGCTCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAG 2205
Db 27879 CTTCTCAGTCTTACTCTCTCTGCTCTGGACAACTGT-CTCCAGATCTGTCACTATCCAAG 27821
Qy 2206 GGGGTCCNTAAGACGGGCACTAGATCTTTTCCAGCCCACTAAGTTATGAATGAGTGG 2265
Db 27820 GG-----TCCTGGGACAGCACTAGATAC-TTCTCCAGCCCACTAAGTGTG-ACTGG 27770
Qy 2266 GGACTTTATTTTTCATCATGCTTTTCTAATATGCTTTGAAAGCCCACTACCTTTGTTA 2325
Db 27769 GGAATTTACTTTTTCATGCTTTTCTAATATGCTTTAATGCTTAAACCCCACTCCCTTTGTTA 27710
Qy 2326 GGGAGAGACATTTAGCAAAAGCAGGGGCCATTTATACACCTGAAACATAGGAGAGGAACA 2385
Db 27709 GGGAAAGACATTTTAAACAAAGCAGGGACCATTTATACACCTGAAACATAGGAAAGGAACA 27650
Qy 2386 CCCGTTTGTGTCCTGCTTGGAGGAGGATTAATCTCCTGAGTCTGGGCAACAGAGG 2445
Db 27649 CCCGTTTGTGT-CCCTGCTTGAAGGAGGATTAATCTTAAAGTCTGGGCAACAGAGG 27591
Qy 2446 ACAATATGAGCAGCAAAAGATGCCCTCTCTGTTCAAGTTTAAACTAAAGGATTTCCACTT 2505
Db 27590 ACAATATGAGCAG-CAAAATATGCTGCTCTGTTTAAACTAAAGGATTTCCACTT 27532
Qy 2506 CTTTCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGA 2565
```



```
RESULT 7
US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15858
; LENGTH: 145320
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15858

Query Match      29.3%; Score 2220.4; DB 4; Length 145320;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2728; Conservative 92; Mismatches 216; Indels 115; Gaps 32;

QY 606 GATGGGAAACGTTCCCGCAGACAAACAAACGCCCCCTAAGACGTATTCCTGGAATTTGGGA 665
DB 97137 GATAGGAAACGTTCCCTCAAGGCAAAACACCCCTAAGATGTATTCCTGGAATTTGGGA 97078

QY 666 MCAATTTGACCTCAGACACTAGAGAGAACGACTTATATCTCTCGAGTGCCTG 725
DB 97077 CCAATTTGACTCTCAGATGCTAAGAAAAAAGAC--ATATTTCTTCGACGTACCCCTG 97020

QY 726 GCA-----CTCCTGAGGAAAGTATAAATATAA 753
DB 97019 GCAACGATATACTTTTAAGGGGAGAAACCTGCGATCCTGAGGAGACATAAATATAA 96960

QY 754 CACATCTTACAGTAGACVCTTTTGTAGAA--AAGGCAATPGAGTGAAGTGCATA 810
DB 96959 CACATCTTACAGTAGACCTCTTTGTAGAAAAGAGGCAAAATGGTGTGAAGTGTATA 96900

QY 811 AGTACAACTTTCTTTTCAATAGAGACAACTCAATATATGTAAAAGTGTGATTTATG 870
DB 96899 CGTACAAACCTTTCTTTTCAATTAAGAGACAACTCGCAATTTATGTAAAAGTGTGATTTATG 96840

QY 871 CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCAGCATCCCC--GACTCCTTCCCC 928
DB 96839 CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTACCCAGCATCCCCCAGACTCCTTCCCC 96780

QY 929 AMYTAATAAGGACCCCCCTTCAACCCCAAAATGGTCCAAAAGAGATAGACAAAAGGGTAAA 988
DB 96779 AAATAATAAGGACCCCCCTTCAACCCCAAAACGGTCCAAAAGAGATAGACAAAAGGGTAAA 96720

QY 989 CAGTGAACCAAGAGTGCATATATCCCAATATGACCCCTCCCAAGCAGTGGGAGAA 1048
DB 96719 CAATAACCAAGAAATGCCAATATATCCCGCATATATGCCCCCTCC--AAGCGTGGGAG 96664

QY 1049 GAGAAATTCGGCCAGCAGAGTGCATGTGCVTTTTTCTCCAGACTTAAAGCAATATAA 1108
DB 96663 GAGAAATTCGGCCAGCAGAGTGCATGTGCVTTTTTCTCTCTCAGACTT----TAAATTA 96608

QY 1109 AAACAGACTTAGGTAATTTCTCAGATAAYCTGTAGTGCATATATGTTGTTTTTAAAGGGT 1168
DB 96607 AAATAGACCTTAGGTAATTTCTCAGATAACCCCTAATGGCTATATGATGTTTTTAAAGGGT 96548

QY 1169 TAGACAACTTTTGTATCTGACATGAGAGATATATATGTCACCTGCTTAATCAGACACTA 1228
DB 96547 TAGACAACTTCTGATCTGATATGAGAGATATA-TATGTTACTGCTTAATCAGACACTA 96489
```



Qy	5652	ATGCGCTGTATGACACGATGAGTCTCCCTCTYACMAGAGTTTCTATGAGAGTAATGACGGCTCC	5711
Db	834	ATGCGCTGTATGACCAAGTAGTACCTCCCTTACCAAGAGTTTCTATGAGAGTAATGACGGCTCC	893
Qy	5712	CGAAATATTGATGCGCCCATCGTATATAGGAGTCTTTSTAAGGGAACCCACACCTTCACTCG	5771
Db	894	CGAAATATTGATGCGCCCATCGTATATAGGAGTCTTTCTAAGGGAACCCACACCTTCACTCG	953
Qy	5772	CCACACCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCGATCAAAATAC	5831
Db	954	CCACACCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCGATCAAAATAC	1013
Qy	5832	TCATATTGACAGGAAATAATCTTAACTGTTGCTCTGGAGGACTTGGAGTCACTGT	5891
Db	1014	TCATATTGACAGGAAATAATCTTAACTGTTGCTCTGGAGGACTTGGAGTCACTGT	1073
Qy	5892	CTGTTGGACTTACTTCACCCAAACTGGTATGCTGATGGGGTGGAGTTTCAGAGTCAAGC	5951
Db	1074	CTGTTGGACTTACTTCACCCAAACTGGTATGCTGATGGGGTGGAGTTTCAGAGTCAAGC	1133
Qy	5952	AAGAGAAAAACTGTAAAGAAATACTCTCCAACTCACCGGGGTACATGGCACTCTAG	6011
Db	1134	AAGAGAAAAACTGTAAAGAAATACTCTCCAACTCACCGGGGTACATGGCACTCTAG	1193
Qy	6012	CCCTCAAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCATCTCGCCT	6071
Db	1194	CCCTCAAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCATCTCGCCT	1253
Qy	6072	GTTAAGCCTATTAAATACCAACCTCACTGGGCTCCATGAGGTTCTGGCCCCAAACCTAC	6131
Db	1254	GTTAAGCCTATTAAATACCAACCTCACTGGGCTCCATGAGGTTCTGGCCCCAAACCTAC	1313
Qy	6132	TAACCTGTTGGATATGCCCTCCCTCGAATTCARGCCATATGTTTCAATCCCTGCTACTGA	6191
Db	1314	TAACCTGTTGGATATGCCCTCCCTCGAATTCARGCCATATGTTTCAATCCCTGCTACTGA	1373
Qy	6192	ACAAATGAAACAACTTCAGCAGAGAAATAAACCACTTCGGTTTATAGTAGGACCTCTTGT	6251
Db	1374	ACAAATGAAACAACTTCAGCAGAGAAATAAACCACTTCGGTTTATAGTAGGACCTCTTGT	1433
Qy	6252	TTCAAATSTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTTAGCAATACTAC	6311
Db	1434	TTCAAATCTGGAAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTTAGCAATACTAC	1493
Qy	6312	ATACACAACCAACTCCCAATGCATCAGTGGTAACCTCCCAACACAAATAGTCTGCCT	6371
Db	1494	ATACACAACCAACTCCCAATGCATCAGTGGGTAACCTCCCAACACAAATAGTCTGCCT	1553
Qy	6372	ACCTCAGGAATATTTTTTGTCTGTGTAACCTCAGCCCTATCGTTGTTGAATGGCTCTTC	6431
Db	1554	ACCTCAGGAATATTTTTTGTCTGTGTAACCTCAGCCCTATCGTTGTTGAATGGCTCTTC	1613
Qy	6432	AGAACTATATGTCCTCTCATTTCTTAGTGCCCCVATGRCATCTTACACTGAAACAGA	6491
Db	1614	AGAACTATATGTCCTCTCATTTCTTAGTGCCCCCTATGACCACTCTACACTGAAACAAGA	1673
Qy	6492	TTTTATACAGTTATGCTATATCTAAGCCCCCGCAACAAAGAGTACCCCATCTCTCTTTTGT	6551
Db	1674	TTTTATACAGTTATGCTATATCTAAGCCCCCGCAACAAAGAGTACCCCATCTCTCTTTTGT	1733
Qy	6552	TATAGGAGCAGGAGTGTAGTGCACTAGGTACTTGGCAATGGCGGTATCAAACTCTAC	6611
Db	1734	TATAGGAGCAGGAGTGTAGTGCACTAGGTACTTGGCAATGGCGGTATCAAACTCTAC	1793
Qy	6612	TCAGTTCTATCAAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTGCGCGACTC	6671
Db	1794	TCAGTTCTATCAAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTGCGCGACTC	1853
Qy	6672	CCTGGTCACTTGCAGATCAACTTACTCCCTAGCAGCAGTACTCTTCTAAATCAAG	6731
Db	1854	CCTGGTCACTTGCAGATCAACTTACTCCCTAGCAGCAGTACTCTTCTAAATCAAG	1913
Qy	6732	AGCTTTAGACTYGGCTAAACCGCTGARAGAGGGGGAACTGTTTATTTTTTAGGGGAAGATG	6791

[illegible]

```

RESULT 9
US-09-573-080A-26
; Sequence 26, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0

```



```

: SEQ ID NO 26
: LENGTH: 8399
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: repeat region
: LOCATION: (1)..(839)
: OTHER INFORMATION: henv9
: PUBLICATION INFORMATION:
: AUTHORS: Jurka, J; Malchiewicz, J; Milosavljevic, A
: TITLE: Prototypic sequences for human repetitive DNA
: JOURNAL: Journal of Molecular Evolution
: VOLUME: 35
: ISSUE: 4
: PAGES: 286-291
: DATE: 1992-10-
: DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
: DATABASE ENTRY DATE: 1996-01-26
: DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-26

Query Match      28.6%; Score 2168; DB 4; Length 8399;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;

Qy      605 TGATGGAAACGTTCCCGCAAGACAAACAGCCCTTAAGACGTATTCTTGGARAATTGGG 664
Db      1276 TGATAGGAACACTCAGGCATCAACCGCTCACCTTGAATGTATCTTAAGCCATTGGG 1335
Qy      665 AMCAATTGACCTTCAGACACTAAGAAAGAAACGACTTATATCTTCTGCAGTCCCGCCT 724
Db      1336 ACCAAATTGACCCACAAACCTGAAAGAGGCGACTCATTTTTTCTGCACACTACGCT 1395
Qy      725 GGCAC-----CCTGAGGGAAGTATAAATTATA 752
Db      1396 GACCTTAATATCTCTCTGTAGGGGAAATAAGCCACCTTGAGGGGAAGTATAAATTACA 1455
Qy      753 ACACCATCTTACAGCTAGACVCTCTTTGT---AGAAAAGGCAATGAGATGAAGTGCAT 809
Db      1456 ATACTATCTCGAGTTGACCTTTCTGTAGAGGGAGGTAAATGAGATGAATACCAT 1515
Qy      810 AAG----TACAACTTTCTTTTCAATTAAGAGACAACCTCACAAATTATGTAAAAAGTGTGAT 865
Db      1516 AAGGTATTACAAGCTTCTTTTCACTGAAGGAGAATACAACTATGCAAGCTTGCAT 1575
Qy      866 TTATGCCCTACAGAAAGCCTTCAGAGTCTACCTCCCTATCCAGCATCCCGGACTCCTTC 925
Db      1576 TTACATCCCAAGGAGACCTTTTCAGCTTACCCCCATATCTTAGACCTCCCTAGAGCTCTC 1635
Qy      926 C--CCAMYTAAAGGACCCCTTTCAACCCAAATGTGTCAAAAGGAGATAGACAAAAGG 983
Db      1636 CTTCCTATTATGCGAAGCTCTCTCCATCTCCCTCGCCAGAAGGAATAAGCAAGAA 1695
Qy      984 GTAACAGTGAACCAAGAGTGCCAAATATTCCCAATATTAGCCCTCCCAAGCAGTGGG 1043
Db      1696 ATCTCAAGAGACCACAAAAACCCCTGGGCTATCGGTTATGTCCCTT-CAAGCTGTAGG 1754
Qy      1044 AGGAAGAGAAATTCGGCCAGCAGATGCGATGTCGCTTTTCTCCAGACTTAAGACA 1103
Db      1755 GGGTGGAGAAATTTGGCCCAACCCAGGTACATGTCCCTTCTCCCTCTGTGATTTAAAGCA 1814
Qy      1104 AATAAAACAGACTTAGTAAATCTCAGATAAACCTGATGGCTATATTGTGTGTTTACA 1163
Db      1815 GATCAAGGCAGACTGGGGGAAGTTTTCAGATGATCTCTGATAGGCACATGATGTCTCCA 1874
Qy      1164 AGGCTTAGGCAATTTCTTGATCTGATGGAGAGATATATGTCACTGCTAAATCAGA 1223
Db      1875 GGGTCTAGGCAACCTTCGATCTCACTTGGAGAGATGT-CATGCTATTGTTAGATCAA 1933
Qy      1224 CACTAACCCCAATAGAGAAAGTGCCACCAATAACTGCAGCGCTGAGGTTTGGGATCTCT 1283
Db      1934 CCTCGGCTTAAAGAAAGAAATGTGCTGTAGCTGCAGCTGAGAGTTTGGAGATACCT 1993

```

```
QY 2360 TACACCTGAATAGGAGAGGAACACCCGTTTGTGTTNCCCTGCTTGAGGAAGGAATT 2419
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3054 TCTACATGATATGGGNAACAGTTACCCATTTGTTGT-CCCTACTTGGAGGGGNAATT 3112
QY 2420 AATCTGTAAGTCTGGGCAACAGAGGAACAATATGGACGAGCCAAAGAAATGCCGTCCTGT 2479
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3113 AACCCCTGNACTCTGGGCAATTTGGAAGGACAATTTGGAAGGGCAAAA-AATGCCACCCAGT 3171
QY 2480 TCAAGTTAACTTAAAGGATTCACACTTCCCTTCCCTACCAAGGCAGTACCCCTCAGACC 2539
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3172 CCAAAATCAGGCTTAAAGACTCCACCACTTTTCTCTATCAAAGGCAATATCCCTTAAGGCC 3231
QY 2540 CAAGGCCCAACAAGGATTTCCAAAAGATTGTTAAGGACTTAAAGGCCCAAGGCTTAGTAAA 2599
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3232 TGAAGCTCTTAAAGGATTAAGGATATTTAAACATTTTAAAGCTCAAGGCTTAGTAAAG 3291
QY 2600 ACCATGCAATAACTCCCTGCAAGTAATCCGTAAGTGTAGGAGGCAAGAAACCCAGTG 2659
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3292 GAAATGCAGCAGTCCCTGCAACACCC------AATCTAGGAGTACAAAACCAAAATG 3344
QY 2660 GACAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGGAGGCCGTTGCTCTTTATA 2719
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3345 GTCACCTGGA-GACTAGTGAAGATCTTAGACTCATCAAT-GAGGCGATTAATCCTCTATA 3402
QY 2720 CCCAGCTGTACCTAGCCCTTATCTGTGMYTTTCCCAATATCCAGAGGAAGCAGAGTGTT 2779
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3403 TCCAGTTGTACCCNACCCTTATACCTGCTTTCTCAANTACCAGAGAGAGAGATGTT 3462
QY 2780 TACASTCCTGGACCTTMMAGGATGCTTTCTTGTGATCCCTGTACATCCTGACTCTCAATT 2839
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3463 CATGGTTCTGGACCTCAAGGATGCTTCTGTTTCCC-CTGCACTCTGACTCCAGTT 3521
QY 2840 CTTGTTGGCTTTGAAGATACTTCAAACCCARCATCTCAACTCAGCTGGACTRTTTACC 2899
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3522 TCTGTTGGCTTTGAGGATCCACAGACCAACACGCTCCCACTTACATGGATGGTCTTGCC 3581
QY 2900 CCAAGGTTTCAGGATAGYCCCCATCTATTGGCCAGGCATTTAGCCCAAAGACTTGAGYCA 2959
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3582 ACAAGGTTTAGGATAGCCCTCACTGTTGGTCAGGCCTGAGGCTGAGGCTTAGGCCA 3641
QY 2960 RTYMTACACTGGACACTCTTGTCTTGTGTAGTGTGATGATTTACTTTTGTGCGCYR 3019
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3642 CTTCTCAAGTCCAGGCACTTTGGTCTTCAGTATGTGATGATTTACTTTTGGCTAGCAG 3701
QY 3020 TTGAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMAATTTCTCGYACCTGTGG 3079
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3702 TTCAGAGCCCTCATGCCAGCGGCTACTCTAGATCTCTTGAACTTTTACCTAATCAAGG 3761
QY 3080 CTACAWGTTTCCAAACSBARGCTCARCTCTGCTCAGCAGGTTAAATACTTAGGRC 3139
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3762 GTACAGGCACTTAGTCAAGGTTGAGCTTTGCTTACAGCAGGCTAAATATCTAGGCT 3821
QY 3140 AARATTTATCCAAAGGCACARGGCCCTCAGTGAGGAAYRYATCCAGCCTATCTGGCTTA 3199
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3822 AATCTTAGCCAGAGGACCAAGGCCCTCAGCAAGGAATGAATACAGCCTATCTAGGCTTA 3881
QY 3200 TCCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGCTTAAGYAGGTTCTGCCG 3259
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3882 TCCTTGCCCTTAAGACATTTAAACACAGTTGACAGGGGTTCTTGAATACCCGGCTTTTGCCG 3941
QY 3260 AAWATGGATTTCCAGGTTGTCRAAATAGCCAGGYCATTAATACASTAATTAAGGAAA 3319
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3942 ACTATGGA-TCCCTTGAATACAGTGAGATAGTACAGGCCCTCCACTCTAATCAAGGAGA 4000
QY 3320 CTCAGAAAGCAATACCCATTTARTAAGATGGAAYMCTGAAGYMAAGTGGCTTTCCAGG 3379
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4001 CCCAGAGGCAATATCTATCTAGTAGAATGTTAACCAGGGGCAAGAAACAGCCTTCAAAA 4060
QY 3380 CCCCTTAAAGAGGCTTTAAACCCAAAGYCCGAGTGTTAAGYTTGCCAARGGGCAAGCTT 3439
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4061 CCTTAAAGCAGGCCCTGTGA---CAAGTTCAGCTTTTAAAGCCTTTCCAC-----AAACTT 4111
QY 3440 TTSITYATAYRTACAGAAAGAAAAACNAGAAAGCTCTGGAGTCTCTTACAGRTCCRAG 3499
```

```
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4112 CTCTTATATGTACAG-AGAGAGCAGGATAGTCTCTGGAGTCTTACTCAGACTCTTG 4170
QY 3500 GGNAGAGCTTGGCAACCVRTGGCRYACCTGASTAAGGAAAATGATGTAGTGGCAAGGTT 3559
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4171 GGCACAAACCCACAAACAGTGGCATACCTAAGTAAGGAATTTGATGTAGTACAAAGGCT 4230
QY 3560 GRCYTCATTTTAYYGGGTAGTGGTCAGTGTAGCAGTYKTATGTATCTGAAAGCAGTTAAAA 3619
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4231 GGCCTCACTGTTTAAAGGTAGTTCAGACAGTGGCCATCTTTCATGTCCAGAGGCTATCAAA 4290
QY 3620 TAATA CAGGAGAGATCTTATCTGTGTGGACATCTCATGAKGTGAAYRGATACTCACTG 3679
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4291 TAATA CAGGAA-AGATCTCACTGTCTGGACTACTCATGTGTGTAATGGCATACTAGTG 4349
QY 3680 CTAAAGGAGACTTGTGGCTGTGACAGAAACVGTTTTACTTTAAATRTCAAGCTCTATTACTTG 3739
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4350 CCAAGGAAATTTATGGCTATCAGACTACCACTTCTTAGATACAGGCACTACTCTCTTG 4409
QY 3740 AARGGCCAGTGTGCACTGTGCACTTTGTGCAACTCTTAAACCCAGYCNCAATTTCTCCAG 3799
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4410 AGGCA CAGTGTCTCAAATATCTACATGTGTGGCCCTCAACCTTGCCACTTTTCTCCAG 4469
QY 3800 ACAATGAAGAAAGATARAAYATACTGTCAAACAARTAAATTTCTCAAAACCTATGCCACT 3859
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4470 AGAATGGGAAACCAATTGACATGACTGCGCAACAAATTTACAGTCCAGACTTATGCCACC 4529
QY 3860 GAGGGACCTTGTAGAGTTCCYTTGACTGATCCYGACCTTCAACTTGTATCTAGTATGGA 3919
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4530 AAGATGATCTCTTAGAAGTCCCTTAGCTAATCTGTACCTT-AACTTATATACCGATGGA 4588
QY 3920 AGTTCTCTTTCTAGAAAAAGCACTTCGAAAAAGYGGGGTATGCACTGGTCACTGATAATGGA 3979
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4589 AGTTCTATTGCAAGAAATGGATATGAAGGCGAGTTGTGACATAGTTAGTGTAGTAAT 4648
QY 3980 ATAYTTGAAAGTATCCCTCACTCCAGGAACTAGTGTCTTAGCTTTCAGAGAACTAATAGCC 4039
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4649 GTACCTGAAAGTAAAGCTCTTCCCCAGGACCAAGCAGTAAACAGAACTAGTGGA 4708
QY 4040 YTCATYKGGCCTCTAGAAATTAGGAGAGRAAAAGGYAAATATATATACAGACTCTRAR 4099
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4709 CTTACCTTGAGCTTAGAATCTGGGAAAGGAAAGAAATAAATGTGTATACAGATAGCAAG 4768
QY 4100 TATCTYACCTTAGTCNTCCATGCCCCTATGMRGCAATATGSPAGAAAGGAAATCTCTAAT 4159
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4769 TATGCTTATCTAATCCCTCATGCCCCTATGCAATATGGAAGAAAGGAGTTCCTAACC 4828
QY 4160 TCYAGRGAAACCTATCAMAACATCAGGAAGCCATTAGGARATTTATTTGCGWGTACAG 4219
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4829 TCCGGGGGTACCCCAATTTAAATGCCAAGAGAGATTATGGAGTTATTGCTGCACTGCAA 4888
QY 4220 AAACCTARAGAGTGGWAGTCTTACACTGCGYGGGTCTATCANAAAGGAAAGRAAGGAA 4279
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4889 AAGCCCAAGGAGGTGGAGTCTTACACTGCCCAAGCCATCAGAAAGGTGAAGGAGAAAG 4948
QY 4280 ATASAGRGAAATGCCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGAGGACCCCTCCA 4339
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4949 GCAGAGGAAACCTCAGGAGATGCTGAGGCCAAATTTGCTGCCAGGTGGATATCTCCA 5008
QY 4340 TTAGAAATGCTTTATTAACCTTCTTAGTATAGGGTAAATCCCTTCCGGGAAACCAAGCCC 4399
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5009 TTAGAAATFACCTATATGGAAGGACCTTTGGAATGGAACAAACCCCTCCAAGAGATTAAAGCCC 5068
QY 4400 CAGTACTCAGCAGGAGAAACAGAAATGGGAACTCCACAGG---CAGTTTTCTCCCTCG 4456
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5069 CAGTATTTCCCAAAATGAACACAGATGGGACTCTCATGGGGCATAGTTTTCTCCCTCA 5128
QY 4457 GGACGTTTAGCCTCTGAAGAAAGGGAATACTTTTTGCTGCTGCAACTATCCAATGGAAATTA 4516
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5129 GGTGGTTTAAACGACAAAGAGAGAAAGGCTACTTATATCCGAGCCAGCAGTGGAAAAATA 5188
QY 4517 CTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCACCATCARATGSCAAA 4576
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

Db 5189 CTTAAAAACCTCCACAAACCTTTTCATATGGGTATTTGAAAAACATCATCAAAATGGCCACA 5248
Qy 4577 TCATTATTTACTGACAGGCTTTTCAAACTATCAAGCARATATKTCAGGCGCTGTGAA 4636
Db 5249 TCTCTATTCAGGCGCAATCTACTCTGGGCTTCCACAGGTAGTCAAACTGTGAG 5308
Qy 4637 KTGTGCCARAAATAATCCCTCCCTCTATCGCCAAAGCTCTTCCAGGARAACAARAACA 4696
Db 5309 GTGTGCCAAAGGAATAATCCCTGGTCCATGTTAAGGCCCAATTTGGGGGAACAAGAATA 5368
Qy 4697 GGCATTTACCTGPARAARACTGCAACTGATTTTACCACAA--GCCAAACCTCAGGG 4754
Db 5369 GGTCACTATCTGAGAGGACTGCGAGTTAGACTTTCAACCATATGGCCTAAATCAAAAGG 5428
Qy 4755 ATTTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGTGTGGCARAGGCTTTCCCTG 4814
Db 5429 ATTTCATATCTTGTGTGTGTGTATACCTTTACAATTTGGATAGAGCTTTCCCTG 5488
Qy 4815 TAGACAGAAAGGCCCAAGGTAATAAAGGCACTAGTTCAATGAATAATTTCCAGATT 4874
Db 5489 CAAGACAGAGAAGACTCAGGAAGTGAATGAAGTCTTAATGATGAATAATTTCTAGATT 5548
Qy 4875 CGGACTTTCCCGAGGCTTACAGATGACAATAGCCCTGCTTTCAGGCGCACAGTAACCA 4934
Db 5549 TGGGCTTTCCCTAAGGCTTACAGAGTGAATGTGTCCAG-TTTTAAAGCCCAATAACTCA 5607
Qy 4935 GCGAGTATCCAGGCTTAGGTATACGATATCACTTACACTCGCCTGAAAGGCCACAGTC 4994
Db 5608 GGGATGTCCAGGCTGTAGGATACAGTATCACTTCACTCGCCCGGAGGCCACAATC 5667
Qy 4995 CTGAGGGAAGGTCCGAGAAATGAATGAAYACTCAAGAGACATCTAAAGAAAGCAACCCA 5054
Db 5668 CTGAGGGAAGGTCAAGAAGGCAAAATGAAGAACTCGAGGCACTTAAGGAACCTAACACA 5727
Qy 5055 GGAACCCACTCATCTGCTGTCTGTGCTGTATAGCCTTAAAGAAATCTGCAACTT 5114
Db 5728 GAA----CATCTCCATGGGCTACTCTTTTGCCCATGGCCTTGTGAGAAATCCGAAATTC 5783
Qy 5115 TCCCAAAAGAGGAGCTTAGCCATACGAAATGCTGTATGGAAGGCCCTTCATACCAA 5174
Db 5784 TCCTCAAAATGGGCTCAGTCCATATGAATGCTGTATGGACAACATTTCTTCACAA 5843
Qy 5175 TGACCTTGTCTTGACCAAGACAG-CCAAATTTAGTTGACAGACATCACTCCTTTAGCCAA 5233
Db 5844 TGACCTCTTCTGATGAGGAAGACAACTTTGTTCAAGATATATCTTCTCGGCAA 5903
Qy 5234 ATATCAACAAGTCTTAAACAATTAAGGAACCTATCTCCTGAGAGGGAAGAAAGAACT 5293
Db 5904 ATATCAACAAGCCCTTAAAGCCCTTACCTGAAGGATGTACAGAGAAAGGGAACAGAGTT 5963
Qy 5294 ATTCCACCCWGTGACATGTTATGATCAAGTCCCTTCTCTCTAATTTCCCATCCCTAGA 5353
Db 5964 GTTTCAACAGGAGATCTAGTGTGGTCAATCTCTCCCTCTACCTTCCCTCATCTATGGA 6023
Qy 5354 TACATCTCTGGAAGGAGCCCTTACCCAGTCAATTTATYACCCCAACTCGGTTAAAGTGGC 5413
Db 6024 CTCTCTGTGGGAAGGTCATCTACTCAATAATCTCTTACCCACACTGCACTGAAGTGGT 6083
Qy 5414 TGAGTGGAGTCTTGGATACATACATCTGAGTCAAAATCTCGATATCTGCAAGGAACC 5473
Db 6084 AGGAGTGAATCTTGGATTCCACACACCCGAGTTAAATTTTGGACATCTCCCTGAGGAACC 6143
Qy 5474 TGAATATCCAGGAG 5487
Db 6144 TGTGAGACCATCAG 6157

```

RESULT 10

US-08-979-847B-87

; Sequence 87, Application US/08979847B

; Patent No. 6582703

; GENERAL INFORMATION:

; APPLICANT: PERRON, HERVE

```

;
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
;
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPH
; THERAPEUTIC PURPOSES
;
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P. O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
;
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2304 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 87:
;
; US-08-979-847B-87
;
; Query Match 24.2%; Score 1837.4; DB 4; Length 2304;
; Best Local Similarity 87.2%; Pred. No. 0;
; Matches 2032; Conservative 119; Mismatches 150; Indels 29; Gaps 15;
;
; Qy 2025 TCCAGCAGCAGGAGCTGAGGCTGCTGGGGCAAGCGCATCCCATGCCATCACCTCACAG 2084
; Db 1 TCCAGCAGCAGGAGCTGAGGCTGCTGGGGCAAGTGCAGGCCCATGCCATCACCT--CAG 58
; Qy 2085 AGCCCTGGGTATGCTTACCATTTGAGGCGCAGGAAGG---TCTCTCTGGACACTGGTG 2141
; Db 59 AGCCCGGGTATGTTTACCATTTGAGCATTGAGAGCCAGGAAGTAACTTCTCTCTGGACACTGGCG 118
; Qy 2142 CGGTCTTCTTAGCTTACTCTTCTGTCGCGGACAACTGCTCCAGATCTGTCACTATTC 2201
; Db 119 CAGCCTTCTCAGTCTTACTTCTCTGTCGCCAGACAAATTTGCTCCAGATCTGTCACTATTC 178
; Qy 2202 TGAGGGGGTCCNTAAGACGCGGAGTCACTAGATACTTTTTCAGGCCCACTAAGTTATGAA 2261
; Db 179 ---GAGGGGTCTTAGACAGCCAGTCACTACATAC--TTCCTCTCAGCCACTAAGTTGTG-A 233
; Qy 2262 CTGGGAGCTTTATTTCTTTTCACTGCTTTTCTAAATATGCTTTGAAAGCCCACTACTT 2321
; Db 234 CTGGGAACTTTTACTCTTTTCACTGCTTTTCTAAATATGCTTTGAAAGCCCACTACTT 293
; Qy 2322 GTTAGGGAGACATCTTAGCAAAAGCAGGGCCATTATACCTGACATAGAGAGG 2381
; Db 294 GTTAGGGAGACATTTTAGCAAAAGCAGGGGCCATTATACCTGACATAGAGAGG 353
; Qy 2382 AACACCCGTTTGTGTGTCNCCCTGCTTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAG 2441

```

Db 354 AATACCAATTGCTGT - CCCCTGCTTGGAGGAATTAATCTGAAGTCGGCAATAG 412  
 Qy 2442 AAGGACAAATATGACAGCAAGAAATGCGCTGCTGTTCAAGTTAACTTAAGGATTC 2501  
 Db 413 AAGGACAAATATGACCAAG - CAAAGAAATGCGCTGCTGTTCAAGTTAACTTAAGGATTC 471  
 Qy 2502 ACTTCTCTTCCCTTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTC 2561  
 Db 472 GCTCTCTTCCCTTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAACAAAGGATTC 531  
 Qy 2562 AAGATTTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCAATGCAATCTCCCTGAGT 2621  
 Db 532 AAGATTTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCAATGCAATCTCCCTGAGT 591  
 Qy 2622 AATTCGTTAGTGGATTGAGGAGCAGAAACCCAGTGGACAGTGGAGGTTAGTGCAG 2681  
 Db 592 ACTCC-----AATTTTAGGAGTAAAGGAAACCCAGGACAGTGGAG - GGTAGTGCAG 643  
 Qy 2682 ATCTCAGGATTAATCAATGAGGCGGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTAT 2741  
 Db 644 ATCTCAGGATTAAT - GAGGCTGTTTTTCTCTTAFAACCCAGCTGTACTAGCCCTTAT 702  
 Qy 2742 ACTGTGMYTTCCCAATACAGAGGAGCAGAGTGGTTTACASTCTCGACCTTMAAGGAT 2801  
 Db 703 ACTCTGCTTCCCTAATACAGAGGAGCAGAGTGGTTTACAGTCTCTGGACCTTAAAGGAT 762  
 Qy 2802 GCTTCTTCTGATCCCTGATACATCTCTGATCTCAATCTTGTGTTGCTTTGAGATACT 2861  
 Db 763 GCTTTTTTCTGATCCCTGATCTCTGATCTCAATCTTGTGTTGCTTTGAGATCTCT 822  
 Qy 2862 TCAACCCCAACCTCAACTGACCTGCTGCTGCTTTTACCCCAAGGTTTCAGGATAGYCCC 2921  
 Db 823 TTGAACCCCAACCTCAACTGACCTGCTGCTGCTTTTACCCCAAGGTTTCAGGATAGYCCC 882  
 Qy 2922 CATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGYCAFTYMTCACTCTGGACACTCT 2981  
 Db 883 CATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGYCAFTYMTCACTCTGGACACTCT 942  
 Qy 2982 GTCTTCTGTAAGTATGATTTACTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041  
 Db 943 GTCTTCTGTAAGTATGATTTACTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002  
 Qy 3042 GCCACCAACGCTCTTMAATTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3101  
 Db 1003 GCCACCAACGCTCTTMAATTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062  
 Qy 3102 GCTCARTCTGCTCAGCAGGTTAAATATCTAGGCTTAATATTTATCCAAAGGACCCARG 3161  
 Db 1063 GCTCGGCTCTGCTCAGCAGGATTAGATCTTAGGCTTAAATTTATCCAAAGGACCCARG 1122  
 Qy 3162 CCTCAGTGAAGYATCCAGCTTATCTGCTTATCTGCTTATCTGCTTATCTGCTTATCTGCT 3221  
 Db 1123 GCTCAGTGAAGYATCCAGCTTATCTGCTTATCTGCTTATCTGCTTATCTGCTTATCTGCT 1182  
 Qy 3222 CAACTAAGGCTTCTTGGCTTAAAGYATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3281  
 Db 1183 CAACTAAGGCTTCTTGGCTTAAAGYATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241  
 Qy 3282 CAAATAGCCAGGCTTAAATATCAATTAATTAAGGAACTCAGAAAGCCCAATACCCATTT 3341  
 Db 1242 CCCAATAGCCAGGCTTAAATATCAATTAATTAAGGAACTCAGAAAGCCCAATACCCATTT 1301  
 Qy 3342 ARTAAGATGGAATCAAGYATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG 3401  
 Db 1302 AGTAAGATGGAACCT---ACAAAGTGGCTTTCCAGGCTTAAAGGCTTAAAGGCTTAAAGG 1356  
 Qy 3402 CAAAGCCAGTGTAAAGYATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG 3461  
 Db 1415 CAAAGCCAGTGTAAAGYATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG 1471  
 Qy 3462 AACAAGYATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 3521

Db 1416 AACAGGAATAGCTCTAGGAGTCTTACGCAAGTCTCAGGATGAGCTTCAACCCCTGCT 1475  
 Qy 3522 AACAGGAATAGCTCTAGGAGTCTTACGCAAGTCTCAGGATGAGCTTCAACCCCTGCT 3581  
 Db 1476 ATACCTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1535  
 Qy 3582 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3641  
 Db 1536 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595  
 Qy 3642 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3701  
 Db 1596 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655  
 Qy 3702 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3761  
 Db 1656 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1715  
 Qy 3762 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3821  
 Db 1716 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775  
 Qy 3822 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3881  
 Db 1776 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835  
 Qy 3882 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3941  
 Db 1836 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1894  
 Qy 3942 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4001  
 Db 1895 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1954  
 Qy 4002 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4061  
 Db 1955 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2014  
 Qy 4062 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4121  
 Db 2015 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2074  
 Qy 4122 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4181  
 Db 2075 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2134  
 Qy 4182 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4241  
 Db 2135 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2194  
 Qy 4242 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4301  
 Db 2195 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2254  
 Qy 4302 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4351  
 Db 2255 GGTGAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2304

RESULT 11  
 US-09-949-016-13840  
 ; Sequence 13840, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13840
; LENGTH: 276687
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(276687)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13840

Query Match          19.9%; Score 1507.4; DB 4; Length 276687;
Best Local Similarity 68.3%; Pred. No. 0;
Matches 2889; Conservative 94; Mismatches 914; Indels 52; Gaps 24;

Qy 697 CGACTTATATCTCTCGCAGTCCCGCTGGCACTCTGAGGAAAGTATAAATTATAACAC 756
Db 27184 CCAATATTTCTCTCTGATGGGAAATATGCGCACTTGAGGAACTACAAATTACATAC 27243

Qy 757 CATCTTACAGTACGTCCTTTTGTAGAA---AAGGCAAAATGGAGTGAAGTGCCTAAGT 813
Db 27244 TATCTGAGCTTGACCTTTTCTGTAAGAGGGAAGGCAAAATGGAGTGAATACCTTATGT 27303

Qy 814 ACAAACTTTCTTTTCAATTAAGAGCAACTCAAAATATGTAATAAGTGTGATTATGCC 873
Db 27304 CCAAGCTTTCTTTTCAITGAGGGAATATACAGCTGTGCAAGCTTGCAATTTTACATCC 27363

Qy 874 TACAGGAAGCTTTCAGAGTCTACCTTCCCTATCCAGCATCCG---CGACTCTCTCCCCAM 930
Db 27364 CACAGGAGGACCTCTCAGCTCACCCCATATCTTAGCTCCCTATAGCTCCCTTCCCTAT 27423

Qy 931 YTAATAAGGACCCCTTCAACCCAAATGGTCCAAAGAGATAGACAAAAGGGTAAACA 990
Db 27424 GAAATGATAATCTCTCTAAATCTCCCGCCGCGGCAAGAAATAGCAAGAAATCTCCA 27483

Qy 991 GTGACCAAAAGAGTCCCAATATTTCCCAATATGACCCCTCCCAAGCAGTGGGAGGAAGA 1050
Db 27484 AAGGACCAAAACCCCTCTGCTATATGGTATATGTTCCCTCTCAAGCTGTAGGGGGAGGG 27542

Qy 1051 GAAITCGGCCAGCCAGAGTGCATGCTGCTTTTCTCCAGACTTAAAGCAATAAAA 1110
Db 27543 GAAITGGCCCAACCCGGGTACATGCTCCCTTCTCCCTCTCTGATTTAAGCAATCAAG 27602

Qy 1111 ACAGACTTATAGTAAATTTCTCAGATAAYCTGATGGCTATATTTGTTTAAAGGGTTA 1170
Db 27603 GCAGACCTGGGGAAGTTTTTCAGATGATCTGATAGGTACATAGATGCTCTACAAGGTCTA 27662

Qy 1171 GGACAAATTTTGTATCTGACATGGAGAGATATATATGTCATGCTGTAATATCAGACACTAAC 1230
Db 27663 GGGCAAACTTTCGACCTCGCTGGAGAGATGT-CATGCTACTGTGAGATCAAACTTGGC 27721

Qy 1231 CCCAAATCAGAGAAGTGCCACCATACCTGAGCTGAGGTTTGGCGATCTCTGTATCT 1290
Db 27722 CTTTAATTAAGAATGCGCTTTAGCTGCGCCCAAGAGTTTGAGATACCTGTGTATCT 27781

Qy 1291 CAGTCAGGTCAATGGATANGGATGACAAACAGAGGAAGAAANAATGATTTCCCAACAGGCCA 1350
Db 27782 TAGTCAAGTAAATG---ATAGAATGACAGCCGGAAGAAAGGACAAATTTCCCTACCGGTCA 27838

Qy 1351 GCARGCAGTCTCCAGTCTASACCTTCATTTGGGACACAGAAATCAGTAACATGGGAGATT 1410
Db 27839 GCAACCCATCCCCAGTATGGATGCCCACTGGGACC---TTGACTTCAGATCATGCGGACT 27894

Qy 1411 GGTGCTGACAGACATTTGCTTAACCTTGCTGCTASAGGACTAAGGAAACTASAGGAAR 1470
Db 27895 GGAGTCATAAACATCTGTTGACCTGTATTTAGAAAGGACTAAGGAGAATTA-GAAAAAAG 27953

Qy 1471 TCTAYGAATTAATCAATGATGTCCACCATAACACAGGGGGAAGGAAATCCCTACTGTC 1530

```

```

Db 27954 CCCATGAATTAATCAATGATGTCGCCCAATAACTCA-GGGAAGGAAGAAAATCCTTCTGTC 28012
Qy 1531 CTTTCTGAGAGACTAAGGGAGGCAATTCAGAAAGCGTGCCTCTCTGTACCTGACTCTTC 1590
Db 28013 CTTCTCGAGCGGCTACGGGAGGCGCTTAAGAAAATATATACTCCCTGTCAACCCGAATCACT 28072
Qy 1591 TGAAGGCCAACTAATCTTAAAGCGTAACTTTATCACTCAGTCAGCTGAGACATTTAGAAA 1650
Db 28073 TGAGGGTCAATTTGATTTCTAANAAGATAGTTTATTACCCAGTCAGCCACAGATATCAGGAG 28132
Qy 1651 AAACCTTCAAAAGTCTGCCGTAGGCGGAGCAAACTTAGAAAACCTTATTGAACCTTGGCA 1710
Db 28133 AAAGCTCGAAAAGCAATCCAGGGCCCTGAAACAAAATCTTAGAGGCAATTTATAAACCTGGCA 28192
Qy 1711 ACYTCGGTTTTTTTATATAGAGATCAGGAGGAGCGGACAGGACAAAGCGGATTTAA 1770
Db 28193 ACCTCGGTGTTTCTATATAATAGGACCAAGAGAAACAGGCCCAAAAGGAAGGAGATCAG 28252
Qy 1771 AAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCAAGTGACTTTTGAGGCGCTCTGGAAA 1830
Db 28253 AGAA-AGGCCGCGCTTAGTCATGGCCCTCAGACAAACAAACCTTGGTGGTTCAGAGAG 28311
Qy 1831 AGGAAAAGCTGGGCAAAATTCGAATGCCCTAATAGGCTTTGCTTCCAGTGGGCTCAAGG 1890
Db 28312 GACAGAAAATGGAGCAGGCCAATCACCTGGTAGGGCTTTGTTATCAGTGTGTTTACTAGG 28371
Qy 1891 ACATCTTAAAAAAGATTTCTCCCAAGTAGAAGTAAAGCCGCCCTTCTGTCATGCCCTTATT 1950
Db 28372 ACATCTTAAAAAAGATTTCTCCCAATGAGAAAACAGGCTGCCCTCTGTCATGTTCCACTATG 28431
Qy 1951 TCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGGACAAAAGGTCTTTTGTAGTCAGAAAG 2010
Db 28432 CCGAGGCAATCACTGGAAGGTGCACTGCCCCAGAGGACAGATTTCCCTGGGTGAGAGG 28491
Qy 2011 CACTAAACAGATGATCAGAGCAGGAGCTAGGAGTGCCTGGGGCAAGCGCAATCCCATGC 2070
Db 28492 CCCCACACAGATGATCAACAAACAGGACTGAGGGTGCCCGGGCAAGGCCAGCTCATGT 28551
Qy 2071 CATCACCTCACAGAGCCCTGGGTATGCTTACCATTTGAGGCGCCAGGAA---GGTTGTCT 2127
Db 28552 CATCACCTCTACTGAGCCCGGGTATGTTAACTATTGAGGGCCAGGAAATTTGACTTCT 28611
Qy 2128 CCTGGAACACTGGTGGGTCTTCTTGTCTTACTCTTCTGTCGGGACAACTGTCTCTCCAG 2187
Db 28612 CCTGGAACACTGGGAGGCTTCTCAGTGTAACTCTCTGTCGATGATGCTCTCAGG 28671
Qy 2188 ATCTGTCACTATCTGAGGGGTCCNTAAGACGGGAGTCACTAGATCTTTTCCCAGC 2247
Db 28672 ----TCCCTTACCATCCGAGGAATCCTGGGACAGCTGTAAACCAGGTA-TTCTCCACC 28726
Qy 2248 CACTAAGTTATGAACCTGGGAGCTTTATCTTTTCACTGCTTTTCTAATTTATGCTTGA 2307
Db 28727 TCCTCAGTTGT-AAATTGGAAGACTTTGCTCTTTTCACTGCTTTTCTTTGTTGCTGTA 28785
Qy 2308 AGCCCCACTACTGTTTAGGGAGAGACATCTAGCAAAAGCAGGGCCATTTATACACCTG 2367
Db 28786 AGTCCCACTACTTTTAGGGAGGNTATATTAGCAAGGTGGAGCTTTATCTACATG 28845
Qy 2368 AACATAGGAGAAAGCAACCCGTTTGTGTNCCCTGCTTGGAGGAAGAAATTAATCTTGA 2427
Db 28846 AATATGAAAACAAAGTTTACCCATTTTGTGT-CCCTACTTTCAGGAGGGAATCAACCTTGA 28904
Qy 2428 AGTCTGGGCAACAGAGGACATATGACGAGCCAAAGAAATGCCGCTCTGTTCAAGTTA 2487
Db 28905 AGTCTGGGCAATTGGAAGGACAAATTTGGAAG-GGCAAAAAATGCTCTGCGCCAGTCCAATCA 28963
Qy 2488 AAATTAAGGATTCACATCTCTCTCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCC 2547
Db 28964 GGTTAAGAAGTCCCACTCTTTCTTCTTATCAAGAGGCAATATCCCTTAAGGCTGAGCTC 29023
Qy 2548 AACAGGATTCAAAAGATTTGTTAAGGACTTTAAAGCCCAAGGCTTAGTAAACCATGCA 2607
Db 29024 ATAAAGGATTACAGAAATATTGTTAAACATTTGAAAGCTCAAGGCTTAGTGAGGAATGCA 29083

```

QY 2608 TAACTCCCTGCAGTAATTCGGTAGTGATGAGGAGGACAGAAACCCAGTGAGCAGTGG 2667  
Db 29084 GCAGTCCCTGCAACACACCC-----AATTCAGAGTACAAAACCTGAACGGTCAGTGG 29136  
QY 2668 AGGTTAGTGCAGAGATCTAGGATTAATCAATGAGGCGGTGTCTTTTATACCCAGCTG 2727  
Db 29137 A-GACTAGTGAAGATCTTAGACTCATCAAT-GAGGCAGTAATTTCCACTATATCCAGTTG 29194  
QY 2728 TACTAGCCCTTATACGTGMYTTCCCAATACAGAGGAGCAGAGTGGTTTACASTCC 2787  
Db 29195 TACCAACCCCTATACCCCTCTCTCAATATACGAGGAGCAGAAATGGTTACGGTTC 29254  
QY 2788 TGGACCTTMAAGGATGCCCTTCTTCGCACTCCCTGTACATCTCTCAATCTTCTTGTG 2847  
Db 29255 TGGACCTCAAGGATGCCCTTCTCTGTATTCCCTGCACCTGTGATTCAGTTCCTCTTTG 29314  
QY 2848 CTTTGAAGATATCTTAAACCCARCATCTCAACTCAGCTGACATCTTACCCCAAGGGT 2907  
Db 29315 CTTTGAAGATATCTTAAACCCARCATCTTCAACTTACCTGACGCTTTCGCCCAAGGGT 29374  
QY 2908 TCAGGATAGYCCCACTCTATTGCGCAGGCACTTAGCCCAAGCTTCGAGYCATYMTCAT 2967  
Db 29375 TTAGGATAGCTCTCATCGGTTTGGTCAGGCCCTTAGCCCAAGATCTAGGCCACTTCTCAA 29434  
QY 2968 ACCTGGACACTTGTCTTCRGTAKGTGGATGATTTTCTTTTRGCGCCYRITTCAGAAA 3027  
Db 29435 GTCCAGGCACTCTGGTCTTCAATATGTGGATGATTTTCTTTGGCTACCACTTCAAG 29494  
QY 3028 CTTTGTGCCATCAAGCCACCCAGCRCTCTTMAATTTCTCGCYVACCTGTGGCTACAWGG 3087  
Db 29495 CTTGTGCCAGCAGCTACTCTAGATCTCTTGAATTTCTAGCTGATCAAGGGTACAAAGG 29554  
QY 3088 TTTTCCA-AAGSARAGTCACTCTGCTGCAGCAGGTAAATTAATCTTAGGCTTAARATTA 3146  
Db 29555 TGCTAGTGAAGGCCCGGCTTGGCTACAGCAGGTAAATTAATCTTAGGCTTAATCTTA 29614  
QY 3147 TCCAAAGGCACACGCGCCCTCAGTGAGGAAYRYATCCAGCCTATCTAGGCTTATCTCAT 3206  
Db 29615 GCCAAAGGACCGGCGCCCTCATCAAGGAGGATACAGCTTACTGGCTTATCTTGC 29674  
QY 3207 CYCAAAACCCCTAAAGCAACTAAGRGRTCTCTGGCTTAAYAGYTTCTGCCGAATATGG 3266  
Db 29675 CTTCAGACATTAACAACAGTTGTGGGGTCTCTTGAATTAATCTAGGCTTGGCCGACTATGG 29734  
QY 3267 ATTCCCGAGTGTGCRAAATAGCCAGYCATTAATATACASTAATTAAGGAATCTCAGAA 3326  
Db 29735 A-TCCCGGATACAGCGAGATAGCCAGGCCACTCTGTACGCTAATCAAGGACACCCAGAG 29793  
QY 3327 AGCCAATACCCATTTATTAAGATGGAYAMCTGAAGYRAAGTGGCTTTCCAGGCGCCCTAA 3386  
Db 29794 GGCAATATAC-----TCTGGTGGATCAGAGCAGAAACAGCTTCAAAACCTTAA 29845  
QY 3387 AGAAGGCTTAAACCCAGYCCCAAGTGTAAAGYTTGCAACRGGGCAAGACTTTTSTTYA 3446  
Db 29846 GCAGGCCCT--AGTACAAGCTCCAGCTTTAAGCCCTTCCACAGTACAAAAGCTTCTCTTA 29903  
QY 3447 TAYTCAAGAAAAAACAAGAAAYAGCTCTGGAGTCTTACACAGRTCCBAGGAGCAG 3506  
Db 29904 TACGTCACAG-AGAGAGCCGGATAGCTCTTGGAGTCTTACTCAGACTCGTGGGACAC 29962  
QY 3507 CTTGCAACCYRTGGCRVACCTGASTAAGGAAAYTATGTAGTGGCAAGGGTTCRCYTCA 3566  
Db 29963 CCCACCCAGTGGCATACCTAAGTAAGAAATTTAGTGTAGCAAAAGGCTGGCTCA 30022  
QY 3567 TTGTTTAYGGTAGTGTGGCAGTAGCAGTYKTATCTCTGAAGCAGTTAAATAATACA 3626  
Db 30023 CTGTTTAAGTGTAGTTCAGCAGTGGCGCTCTTAGTGTACAGACTATCAAAATAATACA 30082  
QY 3627 GGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYGCACTACTCACTGCTAAGG 3686  
Db 30083 AGGAAAGGATCTCATGTCTGACTACTCATGATGTAATGGCACTACTAGGTGGCAAGG 30142

QY 3687 AGACTTGTGGCTGTCCAGACAAACGTTTACTTAAATRTCAGGCTCTATTCTTGAARGGCC 3746  
Db 30143 AAGTTTATGGCTATCAGACAGCCACCTACTTAGATACAGGTCCTCTCTTGAAGGACC 30202  
QY 3747 AGTGCTGCACTGTGCACCTGTGCAACTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGA 3806  
Db 30203 CATGTTTCAATACGACGCTGTGTGGCCCTCAACCCCTGCCACTTTTCTCCAGAGATGG 30262  
QY 3807 AGAAAAAGATARAAYATAAAGTGTCAACAARTAAATTTCTCAAACTATGCCACTCGAGGGGA 3866  
Db 30263 GGAACCAATGACCATGACTGCCCAACAAATTAATAGTCCAGACTTATGCTCCCGAGATGA 30322  
QY 3867 CTTTGTAGAGTTCCTTGTGATCTGATCCYGACCTTCAACTTGTATCTACTGATGGAATTCCT 3926  
Db 30323 TCTCGTAGAGTCCCTTAACTAATCTGACCTT-AACCTATATCTAGTGAAGTTCAT 30381  
QY 3927 TTTAGAAAAAGACTTTCGAAAAAGYGGGTATGCAGTGGTCAGTGAATAATGGAAATVTTG 3986  
Db 30382 TTTGGGAATGGATAGAGGAGGAGGTTACGCCATAGTGTAGTGTAAACCATCTTG 30441  
QY 3987 AAAGTAAATCCCTCACTCCAGGAAGTGTCTTGTAGTGTGCGAGAACTAAAT 4035  
Db 30442 AAAGCAAGCTTTACCCAGAGACCACTGCCAGTTAGTGAAGTGTAGT 30490

## RESULT 12

US-08-979-847B-88

; Sequence 88, Application US/08979847B

; Patent No. 6582703

; GENERAL INFORMATION:

APPLICANT: PERRON, HERVE  
BESEME, FREDERIC  
BEDIN, FREDERIC  
PARANHOS-BACCALA, GLAUCIA  
KOMURIAN-PEADEL, FLORENCE  
JOLIVET-REYNAUD, COLETTE  
MANDRAND, BERNARD  
GARSON, JEREMY  
TUBE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS

ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPH

THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF &amp; BERRIDGE, PLC

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: &lt;unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 39046A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 88:

SEQUENCE CHARACTERISTICS:

LENGTH: 2364 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 88:



US-08-979-847B-88

Query Match 19.2%; Score 1452.4; DB 4; Length 2364;  
Best Local Similarity 79.1%; Pred. No. 0;  
Matches 1813; Conservative 114; Mismatches 266; Indels 99; Gaps 18;

Qy	2021	ATGATCCAGCAGCAGGAGCTGAGGGTGCTGGGGCAAGCGGCATCCCATGCCATCACCCCTC	2080
Db	1	ATGATCCAGCAGCAGGAGCNGAGGGTGCCCGGGGCAAGCGGCATGCCATCACCCCTC	60
Qy	2081	ACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAAAGT-----TGTCCTCTGGACACT	2137
Db	61	ACAGAGCCCCAGGTATGCTTGACCAATTGAGGGGTGAGAAAGGGTNACTGTCTCTGGACACT	120
Qy	2138	GGTGCGGTCTTCTTTAGTCTTACTCTTGTCTCCGGAACAATGTCTCTCCAGATCTGTCACT	2197
Db	121	GGCGGNGCCTTCTCAGTCTTACTTTTCTGTCTGCTGGAACAATGTCTCTCCAGATCTGTCACT	180
Qy	2198	ATTCTGAGGGGGTCNTAAGACGGGAGTCACTAGATCACTTTTTCAGGACCACTAAGTTA	2257
Db	181	GTCC--GAGGGGTCTTAGGACACCACTAGATAC-TTCTCCGAGCACTTAAGTTG	236
Qy	2258	TGAACCTGGGAGCTTTATCTTTTTCACATGCTTTTCTAAATTATGCTTGAAGCCCACTA	2317
Db	237	TG-ACTGGGGAACCTTACTCTTCCACATGCTTTTCTAAATTATGCTTGAAGCCCACTC	2395
Qy	2318	CCTTGTTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATACACTGAAACATAGGAG	2377
Db	296	TCCTTTGGGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATACATGTGAATATAGGAG	355
Qy	2378	AAGGAACACCGGTTTGTGTGTCNCCCTGCTTGAGGAAGGAATTAATCTGAAGTCTGGGCA	2437
Db	356	AAGGAACAACGTGTTGTGTGTG-CCCTGCTTTGAGGAAGGAATTAATCTGAAGTCCGGCA	414
Qy	2438	ACAGAAGGACAAATATGGACGAGCAGCAAGAAATGCCCGTCTGTTCAAGTTAAACTAAAGGA	2497
Db	415	ACAGAAGGACAAATATGGACAAAG-CAAAGAAATGCCCGTCTGTTCAAGTTAAACTAAAGGA	473
Qy	2498	TTCCACTTCTCTTTCCCTTACAAAGGCAAGTACCCCTCAGACCCCAAGGCCCAACAAGGATT	2557
Db	474	TTCCACCTCTCTTTCCCTTACCAAGGCAAGTACCCCTCAGACCCGAGACCCCAACAAGAACT	533
Qy	2558	CCAAAAGATTGTTAAGGACCTTAAAGCCCAAGGCTTAGTAAACCACTGCATACCTCCCTG	2617
Db	534	CCAAAAGATTGTTAAGGACCTTAAAGCCCAAGGCTTAGTAAACCAAGCAATAGCCCTTG	593
Qy	2618	CAGTAATTCCTGTAGTGGATTGAGGAGGCAAGAAACCCAGTGGACAGTGGAGGGTTAGTG	2677
Db	594	CAAGACTCC-----AATTTTAGGAGTAAGGAAACCCACGACAGTGA-GGTTAGTG	645
Qy	2678	CAAGATCTCAGGAATTATCAATGGAGGCGGTGCTCTTTTATACCCAGCTGTACTAGCCC	2737
Db	646	CAAGAACTCAGGAATTATCAAT-GAGGCTGTGTCTCTCTATACCCAGCTGTACTTAACCC	704
Qy	2738	TTATACTGTGMYTTTCCCAATACACAGAGGAAGCAGAGTGGTTTACHSTCTCGACACTTAA	2797
Db	705	TTTATACAGTGCTTTCCCAATACACAGAGGAAGCAGAGTGGTTTACAGTCTCTGACACTTAA	764
Qy	2798	GGATGCTCTTCTGCACTCCCTGTACATCCTGACTCTCAATTCTGTTGGCTTTTGAAGA	2857
Db	765	GGATGCTCTTTCTGCACTCCCTGTAGCTCTGACTCTCAATTCTGTTGGCTTTTGAAGA	824
Qy	2858	TACTTCAAACCCARCATCTCAACTCAGCTGGACTRTTTTACCCCAAGGGTTCAAGGATAG	2917
Db	825	TCTTTTGAACCCCAACGCTCAACTCACTCGACTGTGTTTACCCCAAGGGTTCAAGGATAG	884
Qy	2918	YCCCCACTATTTTGGCCAGGCATTTAGGCCAAGACTTGAGYCARTYMTCACTCTGACAC	2977
Db	885	CCCCCACTATTTGGCCAGGCATTTAGCCCAAGACTTGAGTCAATTTCTATCTGCTGACAC	944
Qy	2978	TCTTGTCTTTCRGTAGTGGATGATTTACTTTTTTRGCGCCYRITTCAGAAACCTTGTGCCA	3037
Db	945	TCTTGTCTTTCAGTACATGATGATTTACTTTTGTAGTCGCCGTTTCAGAAACCTTGTGCCA	1004



Db 1125 CAGGGCCCTCAGTGAAGGAAGTATCCAGCCTATATCTGGCTTATCTCTCATCCCAAAACCCT 1184  
Qy 3218 AAGCAACTAAGRRRTCTTGGCTTAAYAGGYTTCTGCGAATATGATTCGCCAGGT 3277  
Db 1185 AAGCAACTAAGAGGGTCTTGGCTTAACAGAGTTCTGCGAATATGATTCGCCAGGT 1243  
Qy 3278 WTGCGRAATAGCAGGAGGATTAATATACATTAATTAAGAACTCAGAAAGCCATCCC 3337  
Db 1244 ACASCCCAATAGCAGACCAATATATACATTAATTAAGAACTCAGAAAGCCATCCC 1303  
Qy 3338 ATTATTAAGATGAGYAMCTGAAGYMAAGTGGCTTTCAGGCCCCCTAAAGAGGCTTA 3397  
Db 1304 ATTATTAAGATGAGACCT---ACAGAGTGGCTTTCAGGCCCCCTAAAGAGGCTTA 1358  
Qy 3398 AACCAAGYCCAGTGTAAAGYTGCAACRGGGCAAGACTTTTSTYATAYRTACAGA 3457  
Db 1359 AACCAAGYCCAGTGTAAAGYTGCAACRGGGCAAGACTTTTSTYATAYRTACAGA 1417  
Qy 3458 AAAAACAAGYAGCTCTGGAGTCTTACAGRTCCBAGGAGGAGCTTGCACCCYR 3517  
Db 1418 AAAAACAAGYAGCTCTGGAGTCTTACAGRTCCBAGGAGGAGCTTGCACCCYR 1477  
Qy 3518 TGCCRYACCTGAGTAAGGAAATGATGTAGTGCAAGGGTTCGCTCTCATTTGTTAYGG 3577  
Db 1478 TGGTATCTGAGTAAGGAAATGATGTAGTGCAAGGGTTCGCTCTCATTTGTTAYGG 1537  
Qy 3578 TAGTGTGGCAGTACAGTGTATCTGAAGCAGTATAAATAATACAGGGRAGAGATC 3637  
Db 1538 TAATGGNGCAGTACAGTGTATCTGAAGCAGTATAAATAATACAGGGRAGAGATC 1597  
Qy 3638 TTACTGTGGACATCTCATGAGTGAAYRGATCTACTGCTAAAGGAGACTTGTGGC 3697  
Db 1598 TTACTGTGGACATCTCATGAGTGAAYRGATCTACTGCTAAAGGAGACTTGTGGC 1657  
Qy 3698 TGTACAGCAACGTGTATTAATATCAGCTCTATTCTTGAAGGCCAGTCTGCTCRAC 3757  
Db 1658 TGTACAGCAACGTGTATTAATATCAGCTCTATTCTTGAAGGCCAGTCTGCTCRAC 1717  
Qy 3758 TGTGCACTTGTGCACTTTAAGCCAGYCNCAATTTCTTCCAGCAATGAAGAAAGATAR 3817  
Db 1718 TGTGCACTTGTGCACTTTAAGCCAGYCNCAATTTCTTCCAGCAATGAAGAAAGATAR 1742  
Qy 3818 AAYATACTGTCAACAARTATTTCTCAACCTATGCCACTGAGGGGACCTTTTAGARG 3877  
Db 1743 -----CAAACTTATGCTGCCAGAGGATCTTTTAGAG 1776  
Qy 3878 TTCCYTTGACTGATCCYACCTTCACT---TGTATCTGATGGAAGTTCTTTAGAA 3934  
Db 1777 GTCCCTTAGCCAACTGATCTCACTATATATATCTGATGGAAGTTCTTTAGAA 1836  
Qy 3935 AAGGACTTCGAAAGYGGGGTATGC---AGTGGTCAATGATAATGGAATATTTGAACT 3991  
Db 1837 AAGGATTTACAAAGGAGGATATNCCATAGTGTAGTATGATAAGCAGTACTTGAAGT 1896  
Qy 3992 AATCCCT---CACTCCAGAACTAGTGTCTGATGCGAGAACTAATAGCCCT---CAAT 4045  
Db 1897 AAGCCCTTCTCCCTCCAGGAGCAGCGCCCGCTTAGCAGAACTAGTGGCACTGACCCG 1956  
Qy 4046 KGGGCACTAGATTAGGAGAGGAAAGGGAATATATATACAGACTTATATATGCT 4105  
Db 1957 CAGGCTTTAGAACTTTGGAAGAGGAGGAGGATAAATGTGTATACAGATGCAAGTATGCT 2016  
Qy 4106 YACCTAGTCTNTCC-----ATGCCCATGMRCAATATGS 4138  
Db 2017 TATCTAATCCGAATGCCCATGTTGTTTATCTAATCCGAATATGCCATGTTCAATATGG 2076  
Qy 4139 APAGAAAGGAATTCCTAATCTCYAGAGGAAACCTTATCAMAATCAGGAAGCCATTAGG 4198  
Db 2077 AAGAAAGAGGAGTCTCACTTCTGCGGGAACCCCTTAAATACCAAGTATATCATG 2136  
Qy 4199 AATATTAATYAGGAGTACAGAAACCTTAGAGAGGTGGAGTCTTATACATGCTGGGGTCA 4258  
Db 2137 GAGTTATTTGACACAGTGTCAAAACCTCAAGAGGAGTGGAGTCTTACTGCTCCAAAGCCAT 2196

Qy 4259 CANAAAGGAAGRAAGGAAATATASAAAGRAAYTGCAA 4297  
Db 2197 CAGAAAGGAAGGGAAGAGCAGCATAGTGCTA 2235

RESULT 14

US-09-374-766-57  
; Sequence 57, Application US/09374766  
; Patent No. 6579526  
; GENERAL INFORMATION:  
; APPLICANT: Herve PERRON  
; APPLICANT: Frederic BESEME  
; APPLICANT: Frederic BEDIN  
; APPLICANT: Glauca PARANHOS-BACCALA  
; APPLICANT: Florence KOMURIAN-PRADEL  
; APPLICANT: Colette JOLIVET  
; APPLICANT: Bernard MANDRAND

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL  
; TITLE OF INVENTION: THERAPEUTIC PURPOSES  
; NUMBER OF SEQUENCES: 92  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Oliff & Berridge  
; STREET: 700 South Washington Street, Suite 300  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22314

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/374,766  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/691,563  
; FILING DATE: 02-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berridge, William P.  
; REGISTRATION NUMBER: 30,024  
; REFERENCE/DOCKET NUMBER: WPB 38588  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-836-6400  
; TELEFAX: 703-836-2787  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2391 base pairs  
; TYPE: nucleotide  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-09-374-766-57

Query Match 18.7%; Score 1415.8; DB 4; Length 2391;  
Best Local Similarity 78.4%; Pred. No. 0;  
Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;

Qy 2021 ATGATCCAGCAGCAGCAGTGGGTGCTGGGGCAAGGCCATCCCATGCCATCACCCCTC 2080  
Db 1 ATGATCCAGCAGCAGCAGTGGGTGCTGGGGCAAGGCCATCCCATGCCATCACCCCTC 60  
Qy 2081 ACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGGT---TGTCTCTGGACACT 2137  
Db 61 ACAGAGCCCTGGGTATGCTTGACCATTTGAGGGTCAGAAAGGTNACTGTCTCTGGACACT 120  
Qy 2138 GGTGGGTCTTTAGTCTTACTTCTGTCCTCCGACAACTGCTCCAGATCTGTCACT 2197  
Db 121 GCGGNGCCTTCTCAGTCTTACTTTCTGTCCTGACAACTGCTCCAGATCTGTCACT 180  
Qy 2198 ATTCTGAGGGGTCCNTAAGAGCGGCACTACTAGATACTTTTCTCCAGCCCAAGTTA 2257

No + p r r r



Sequence 53, Application US/08979847B  
Patent No. 6582703  
GENERAL INFORMATION:  
APPLICANT: PERRON, HERVE  
BESEME, FREDERIC  
BEDIN, FREDERIC  
PARANHOS-BACCALA, GLAUCIA  
KOMURIAN-PRADEL, FLORENCE  
JOLIVET-REYNAUD, COLETTE  
MANDRAND, BERNARD  
GARSON, JEREMY  
TUBE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL  
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/979,847B  
APPLICATION NUMBER: US/08/979,847B  
FILING DATE: 26-No. 6582703-1997  
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6400  
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2391 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 53:

US-08-979-847B-53

Query Match 18.7%; Score 1415.8; DB 4; Length 2391;  
Best Local Similarity 78.4%; Pred. No. 0;  
Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;

Qy	2021	ATGATCCAGCAGCAGGACTGAGGTCCTGGGGCAAGCGCCATCCCATGCCATCACCCCTC	2080
Db	1	ATGATCCAGCAGCAGGACGAGGTCCTGGGGCAAGCGCCATCCCATGCCATCACCCCTC	60
Qy	2081	ACAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAGGT---TGTCCTCTGGACACT	2137
Db	61	ACAGAGCCCTGGGTATGCTTGACCAATGAGGGTCAGAGGGGTNACTGTCTCTGGACACT	120
Qy	2138	GGTGGGTCCTTAGTCTTACTCTTCTGTCCTGGGCAACATGCTCTCCAGATCTGTCACT	2197
Db	121	GGCGGCGCTTCTAGTCTTACTCTTCTGTCCTGGGCAACATGCTCTCCAGATCTGTCACT	180
Qy	2198	ATTCGAGGGGTCCTTAAGCAGGCGCAGTCACTAGATACCTTTTCCCGAGCCACTAAGTTA	2257
Db	181	GTCC---GAGGGTCCTAGGACAGCCAGTCACTAGATAC-TTCCTCCAGCCACTAAGTTG	236
Qy	2258	TGAAGTGGGAGCTTTATCTTTTTCATGCTTTTCAATATGCTTGAAGAGCCCACTA	2317
Db	237	TG-ACGGGGAACCTTTACTCTTCCCATGCTTTTCTAATATGCTGCTGAAGAGCCCACTC	295

Qy	2318	CCTTGTAGGAGAGACATCTTAGCAAAAGAGGGGCCATTATACACCTGAACATAGGAG	2377
Db	296	TCTTGTGGGAGAGACATCTTAGCAAAAGAGGGGCCATTATACATGTGAATATAGGAG	355
Qy	2378	AAGGAACACCCGTTTGTGTGTCCTGCTTGAGGAAGGAATTAATCTCAAGTCTCGGCA	2437
Db	356	AAGGAACACCTGTTTGTGTGTCCTGCTTGAGGAAGGAATTAATCTCAAGTCTCGGCA	414
Qy	2438	ACAGAGGACACATATGAGGAGGCAAAAGATGCTGCTCTCTGTTCAAGTTAAAGGA	2497
Db	415	ACAGAGGACACATATGAGGCAAG-CAAGAATGCTGCTCTCTGTTCAAGTTAAAGGA	473
Qy	2498	TTCCACTTCTCTTCCCTACCAAGCAGTACCCCTCAGAGCCCAAGGCCCAACAGGATT	2557
Db	474	TTCCACTTCTCTTCCCTACCAAGCAGTACCCCTCAGAGCCCAAGGCCCAACAGGATT	533
Qy	2558	CAAAAGATGTTTAAAGCATTTAAAGCCCAAGGCTTAGTAAACCATGATCACTCCCTG	2617
Db	534	CAAAAGATGTTTAAAGCAGCTTAAAGCCCAAGGCTTAGTAAACCATGATCACTCCCTG	593
Qy	2618	CAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGGACAGTGGAGGGTTAGTG	2677
Db	594	CAAGACTCC-----AATTTTAGGAGTAAGGAAACCCCAAGCAGTGGGA-GGTTAGTG	645
Qy	2678	CAAGATCTCAGGATTTATCAATGGAGGCGTTGTCCTTTTATACCCAGCTGTACCTAGCC	2737
Db	646	CAAGACTCAGGATTTATCAAT-GAGGCTGTTGTTCTCTATACCCAGCTGTACCTAAGCC	704
Qy	2738	TTATACCTGTGMYTTCCCAATACCAAGAGGAGCAGAGTGGTTTACATCTCGACCTTMA	2797
Db	705	TTATACAGTGTCTTCCCAATACCAAGAGGAGCAGAGTGGTTTACATCTCGACCTTMA	764
Qy	2798	GGATGCTTCTCTGTCATCCCTGTACATCTCTCAATCTTCTGTTGCTTTGAAGA	2857
Db	765	GGATGCTTCTCTGTCATCCCTGTACATCTCTCAATCTTCTGTTGCTTTGAAGA	824
Qy	2858	TACTTCAAAACCCAGCATCTCAACTCACCTGACATCTTTTACCCAGGTTTCCAGGATAG	2917
Db	825	TCTTTGNAACCCAGGCTCAACTCACCTGACATCTTTTACCCAGGTTTCCAGGATAG	884
Qy	2918	YCCCATCTATTTGGCCAGCATTTAGCCCAAGACTTGAGYCATYMTCATCTCGACAC	2977
Db	885	CCCCCATCTATTTGGCCAGCATTTAGCCCAAGACTTGAGTCAATCTCATCTCGACAC	944
Qy	2978	TCTTGTCTTCRGTAAGTGGATGATTTACTTTTGTGCTGCTTTCAGAAAACCTTGTGCA	3037
Db	945	TCTTGTCTTCRGTAAGTGGATGATTTACTTTTGTGCTGCTTTCAGAAAACCTTGTGCA	1004
Qy	3038	TCAAGCCACCAAGCCTCTTTMAATTTCTCGYACCTGTGCTGCTACAGGTTTCCAAAGC	3097
Db	1005	TCAAGCCACCAAGCCTCTTTMAATTTCTCGYACCTGTGCTGCTACAGGTTTCCAAAGC	1064
Qy	3098	ARAGCTCARTCTGTCTCACAGCAGGTTAAATTAATCTAGGCTAAATATATCAAGGAC	3157
Db	1065	AAAGGCTCGGCTGTCTCACAGCAGGTTAAATTAATCTAGGCTAAATATATCAAGGAC	1124
Qy	3158	CARGGCTCTGATGAGAAATATCCAGCTTATCTGCTTATCTCTATCTATCTATCTATCT	3217
Db	1125	CARGGCTCTGATGAGAAATATCCAGCTTATCTGCTTATCTCTATCTATCTATCTATCT	1184
Qy	3218	AAAGCAACTAAGGRRRTTCTTGGCTAAAGGTTTCTGCGCAATGATGATTCCTCAGGT	3277
Db	1185	AAAGCAACTAAGGRRRTTCTTGGCTAAAGGTTTCTGCGCAATGATGATTCCTCAGGT	1243
Qy	3278	WTGCGAAATAGCCAGGCTTAAATATCAATTAAGGAACTCAGAAAAGCCAAATACCC	3337
Db	1244	ACASCCCAATAGCCAGGCTTAAATATCAATTAAGGAACTCAGAAAAGCCAAATACCC	1303
Qy	3338	ATTATTAAGATGAGYATGAGGTTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC	3397
Db	1304	ATTATTAAGATGAGYATGAGGTTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCC	1358
Qy	3398	AACCCAGYCCAGTGTAAAGTTTGCCCAACGAGGCTTTTSTTTATATATATATATATAT	3457

[illegible]